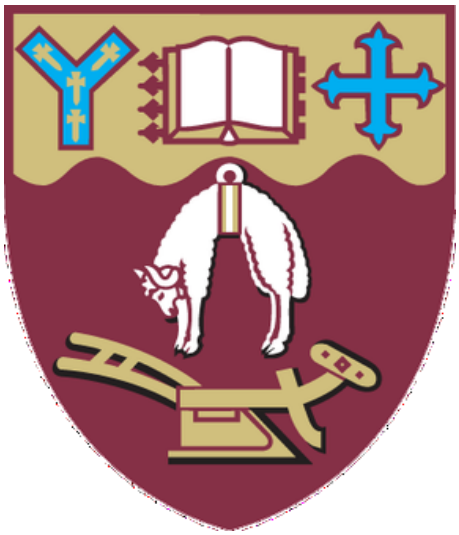
Optimisation of the treatment quality in head-and-neck radiation oncology



Nicholas John Lowther[1,2]

[1]*School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand*

[2]*Department of Radiation Oncology, Wellington Blood and Cancer Centre, Capital & Coast DHB, Wellington, New Zealand*

This dissertation is submitted for the degree of

Doctor of Philosophy (PhD)

January 2020
(Updated May 2020)

DECLARATION

I hereby declare that except where specific reference is made to the work of others, the contents in this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.

Nicholas John Lowther

Originally submitted January 2020. Updated dissertation submitted May 2020

ABSTRACT

Emerging literature has reported reduced treatment toxicity in head-and-neck radiotherapy (HNRT) with 3 instead of 5 mm planning target volume (PTV) margins. However, the loco-regional (LR) control rate was not preserved in all studies. As it was considered whether it was possible to implement reduced treatment margins at the Wellington Blood & Cancer Centre (WBCC) with the aim to improve patients' treatment-related toxicity, it was recognised that many department-specific aspects of HNRT can influence treatment outcomes and should all be considered when the treatment margins are amended. PTV margins are applied to target volumes during treatment planning to account for uncertainties such as patient positioning, geometrical accuracy of the treatment machine and geometrical accuracy of target volume definition. However, these margins do not standardly account for non-rigid anatomy changes (e.g., changes in patient pose, weight loss, tumour response) which are commonly observed during HNRT and can undermine the planned dose objectives. Under the assumption that a loss in target coverage during treatment may occur more often and may become more relevant with reduced mm PTV margins, it was proposed that a safe PTV margin reduction could be achieved by accounting for systematic changes in patient anatomy using timely and appropriate treatment adaption. This approach required quantification and separation of the different modes of anatomical change during treatment, and subsequent investigation of their dosimetric impact. Therefore, the retrospective studies included in this dissertation first investigated the application of deformable image registration (DIR) in combination with Exponentially Weighted Moving Average (EWMA) Statistical Process Control (SPC) charts, and DIR-facilitated dose accumulation. In the first study, DIR between the computed tomography for treatment planning (pCT) images of twelve patients and their daily on-treatment cone beam computed tomography (CBCT) images quantified anatomical changes during treatment to investigate corresponding trends using EWMA charts. The application of EWMA SPC charts showed

that trends in patient positioning of bony anatomy with respect to the first five treatment fractions could only be confirmed at a 90% confidence level in a small number of cases when EWMA process limits were used, whereas absolute patient position deviations could be confirmed in the vast majority of cases when an *a priori* 2 mm clinical limit was used. EWMA process limits were however effective when detecting trends of soft tissue structures. Structure-specific action thresholds for trend detection using SPC charts enabled detection of systematic anatomical changes. The second study defined the intended clinical workflow based on DIR-facilitated dose accumulation to assess the actually delivered dose and the uncertainty in the delivered dose was determined using *in silico* deformations based on clinically observed anatomical changes as ground truth. The uncertainty in DIR-facilitated dose accumulation was accurately quantified and the methodology how to incorporate these prospectively in dose-volume histograms (DVHs) was described. These results demonstrated that the intended clinical workflow is sufficiently accurate to assess the adequacy of target coverage during HNRT. In the third study, the estimated uncertainty derived in the second study was included in the dose reconstruction and accumulation over all fractions for the twelve patients to investigate the robustness of volumetric modulated arc therapy (VMAT) plans that were optimised using either 3 or 5 mm PTV and planning risk volume (PRV) margins. It was shown that loss in target coverage was independent of margin expansion and very patient specific. In addition, it was found that the tightness of target volume coverage at planning was a common factor leading to underdosage. The developed clinical workflow to reconstruct the delivered dose using DIR-facilitated dose accumulation found that PTV/PRV margin reduction did not significantly reduce the robustness of treatment plans to attain adequate target coverage during treatment. This indicates that a safe PTV margin reduction from 5 to 3 mm in HNRT can be achieved. Patient specific verification of the delivered dose using the developed methodology is recommended irrespective of the applied margin expansions considering the patient specific nature of the potential loss in target coverage.

CO-AUTHORSHIP AND THESIS WITH PUBLICATIONS

This serves to explicitly inform examiners that the following dissertation is submitted according to the "*Thesis with publications*" format sanctioned by the University of Canterbury.

If required, more information on this submission format can be found at: https://www.canterbury.ac.nz/media/documents/postgraduate-/information-sheets/Thesis-with-publications-Guidelines.pdf

The main body in this dissertation contains three journal articles. At the time of this dissertation's submission, two of the journal articles have been published/in-press and the third is currently in peer-review. The three articles are presented "*as published/submitted*" except for Figure, Table and Supplementary Data identifiers which have been modified to assist with the readability of this dissertation. In addition, minor formatting changes have been made. Citations of the studies are provided immediately prior to their individual chapters with accompanying text describing the contribution of each author.

The University of Canterbury co-authorship forms which are required to accompany "*any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication*" can be found on the following three pages.

Throughout my PhD candidature I was based within the Medical Physics Department of the Wellington Blood & Cancer Centre, Wellington Hospital as an extra-mural PhD candidate with the University of Canterbury. Dr Robert Louwe took the role of my clinical research supervisor and was also based at the Wellington Blood & Cancer Centre. Dr Steven Marsh took the role of my academic research supervisor and was based at the University of Canterbury.

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter Two in this dissertation is a journal article with the following citation,

Lowther, N.J., Hamilton, D.A., Kim, H., Evans, J.M., Marsh, S.H. and Louwe, R.J., 2019. Monitoring anatomical changes of individual patients using statistical process control during head-and-neck radiotherapy. *Physics and Imaging in Radiation Oncology*, 9, pp.21-27

Please detail the nature and extent (%) of contribution by the candidate:

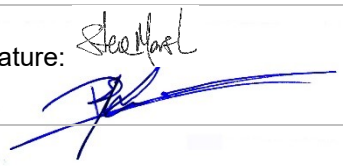
I, along with co-author RL developed the strategy of monitoring individual patients' anatomical changes with SPC and DIR. I collected all patient data, performed all tasks within the treatment planning system and wrote all the MATLAB scripts for data analysis. I integrated standardised uncertainty equations with the mathematics of SPC and performed all the data and statistical analyses. Co-authors DH, HK and JE are consultant radiation oncologists at the Wellington Blood & Cancer Centre and contributed to this study by providing re-contouring of the PGs on CBCTs. I, along with co-authors SM and RL developed the conclusions. I wrote this paper with modifications to structure, flow and readability from SM and RL.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certify that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Steven Marsh* *Rob Louwe*	Signature: 	Date: *17 Dec 2019* *08 Jan 2020*

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:
Chapter Three in this dissertation is a journal article with the following citation,

Lowther, N.J., Marsh, S.H. and Louwe, R.J., 2020. Quantifying the dose accumulation uncertainty after deformable image registration in head-and-neck radiotherapy. *Radiotherapy and Oncology, 143,* pp.117-125.

Please detail the nature and extent (%) of contribution by the candidate:

I developed the intended clinical workflow and the methodology to quantify the dose accumulation uncertainty with the *in silico* ground truth. I collected all patient data, performed all tasks within the treatment planning system, wrote all the MATLAB scripts for data analysis and performed all tasks within Slicer. I performed all the mathematics for the DVH uncertainty and statistical analyses. I, along with co-authors SM and RL developed the conclusions and refined the methods. I wrote this paper with modifications to structure, flow and readability from SM and RL.

Certification by Co-authors:
If there is more than one co-author then a single co-author can sign on behalf of all
The undersigned certify that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: Steven Marsh	Signature:	Date: *17 Dec 2019*
Rob Louwe		*08 Jan 2020*

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:
Chapter Four in this dissertation is a journal article with the following citation,

Lowther, N.J., Marsh, S.H. and Louwe, R.J., 2020. Dose accumulation to assess the validity of treatment plans with reduced margins in radiotherapy of head and neck cancer. In publication in *Physics and Imaging in Radiation Oncology*

Please detail the nature and extent (%) of contribution by the candidate:

I applied the methodology that I developed in *Chapter Three* to simulate the intended clinical workflow for HNRT at the WBCC. I collected all patient data, performed all tasks within the treatment planning system, wrote all the MATLAB scripts for data analysis and performed all tasks within Slicer. I performed all the mathematics for the statistical analyses. I, along with co-authors SM and RL developed the conclusions and refined the methods. I wrote this paper with modifications to structure, flow and readability from SM and RL.

Certification by Co-authors:
If there is more than one co-author then a single co-author can sign on behalf of all
The undersigned certify that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Steven Marsh*	Signature:	Date: *17 Dec 2019*
Rob Louwe		*08 Jan 2020*

ACKNOWLEDGEMENTS

I would like to thank and express my upmost gratitude to supervisors Dr Rob Louwe and Dr Steven Marsh. Firstly, Dr Louwe who has been the force behind improving the quality of patients' head-and-neck radiotherapy treatment at the Wellington Blood & Cancer Centre. I could not have completed, or even started my research if Dr Louwe had not first laid the foundations. I am extremely grateful that Dr Louwe introduced me to the project and continued to persevere through my confused conversations, analyses and attempts to explain myself over the course of the entire PhD work. Secondly, Dr Marsh who provided on-going support, insight and ideas that were of great addition to this research. Furthermore I'd like to thank Steve for addressing my never-ending requests for funding reference letters.

I would also like to thank Ms Lynne Greig and the entire medical physics team at the Wellington Blood and Cancer Centre. Despite not being an employee, the medical physics team consistently provided their time and resources to assist my questions and point me in the right direction when needed. Ms Greig was instrumental in obtaining funding for conference travel – thank you!

The input of David Hamilton, Han Kim and Jamie Evans who are consultant radiation oncologists at the Wellington Blood & Cancer Centre was greatly appreciated. I express thanks for their contouring reviews. In addition I would like to thank Dr Hamilton for his assessment of the potential clinical impact of the results.

This research would not have been possible without the generous funding provided by the Department of Internal Affairs in the form of a Lottery Health Research PhD scholarship and by Universities New Zealand in the form of an Edward and Isabel Kidson scholarship. My appreciation of their funding cannot be understated.

EXAMINERS

I would also like to thank and express my upmost gratitude to my dissertation's examiners Prof Marianne Aznar and Prof Kristy Brock. Their valuable time spent reviewing my dissertation with subsequent commentary enabled me to further consolidate my understanding of this research.

Thank you!

Professor Marianne C. Aznar
School of Medical Sciences
Division of Cancer Sciences
Manchester University
Manchester, UK

Professor Kristy K. Brock
Department of Imaging Physics
Department of Radiation Physics
Director, Program for Image Guided Cancer Therapy Research
University of Texas MD Anderson Cancer Center

ACADEMIC CONTRIBUTIONS

This research has disseminated outcomes according to the following:

- **Department seminar:** "*Progressing Treatment Quality in Head and Neck Radiation Oncology*" for Medical Physics, Department of Radiation Oncology, Wellington Blood & Cancer Centre on the 12th September 2016. PhD project overview

- **Department seminar:** "*Progressing Treatment Quality in Head and Neck Radiation Oncology*" for Medical Physics, Department of Radiation Oncology, Wellington Blood & Cancer Centre on the 31st July 2017. Preliminary results of *Chapter Two*

- **Conference oral presentation:** "*Deformable image registration of planning CT-to-CBCT for quantifying patient-specific anatomical change during head-and-neck radiotherapy in the context of reduced PTV margins*" at the 2017 Engineering & Physical Sciences in Medicine Conference (EPSM), 1st November 2017. Hobart, Australia. This is the premier annual event for medical physicists and biomedical engineers in Australia and New Zealand held annually

- **Conference oral presentation:** "*Quantifying anatomical changes using deformable image registration during head-and-neck radiotherapy to facilitate safe PTV margin reduction*" at the 2017 New Zealand Physics and Engineering in Medicine conference (NZPEM), 22nd November 2017. Palmerston North, New Zealand. This is the annual event for medical physicists and biomedical engineers in New Zealand

- **Department seminar:** "*The efficiency of various 'plan of the day' approaches as a safety net accompanying PTV margin reduction*" for Medical Physics, Department of Radiation Oncology, Wellington Blood & Cancer Centre on the 7th May 2018. Preliminary results of *Chapters Three* and *Four*

- **Department seminar:** "*Deformable image registration (DIR) for dose warping*" for Medical Physics, Department of Radiation Oncology, Wellington Blood & Cancer Centre on the 13th

August 2018. Reporting outcomes from an Australasian workshop and their application to *Chapters Three* and *Four*

- **Conference oral presentation:** "*Daily dose accumulation for individual patients to assess the validity of reduced PTV margin plans in head and neck radiation oncology*" at the 2018 Engineering & Physical Sciences in Medicine Conference (EPSM), 31st October 2018. Adelaide, Australia. This is the premier annual event for medical physicists and biomedical engineers in Australia and New Zealand held annually

- **Published journal article:** Lowther, N.J., Hamilton, D.A., Kim, H., Evans, J.M., Marsh, S.H. and Louwe, R.J., 2019. Monitoring anatomical changes of individual patients using statistical process control during head-and-neck radiotherapy. *Physics and Imaging in Radiation Oncology, 9*, pp.21-27

- **Conference oral presentation:** "*Quantifying the dose accumulation uncertainty after deformable image registration to facilitate safe PTV margin reduction in head-and-neck radiotherapy*" at the 2019 New Zealand Physics and Engineering in Medicine conference (NZPEM), 1st April 2019. Auckland, New Zealand. This is the annual event for medical physicists and biomedical engineers in New Zealand

- **e-Poster:** "*Statistical process control to monitor anatomical changes during head-and-neck radiotherapy*" at the 2019 European Society for Radiotherapy & Oncology (ESTRO) conference held from the 26th – 30th April 2019 in Milan, Italy. This is the premier annual event for European professionals in the radiation oncology field

- **e-Poster:** "*Dose accumulation assessing the validity of reduced PTV margins in head-and-neck radiotherapy*" at the 2019 European Society for Radiotherapy & Oncology (ESTRO) conference held from the 26th – 30th April 2019 in Milan, Italy. This is the premier annual event for European professionals in the radiation oncology field

- **Published journal article:** Lowther, N.J., Marsh, S.H. and Louwe, R.J., 2020. Quantifying the dose accumulation uncertainty after deformable image registration in head-and-neck radiotherapy. *Radiotherapy and Oncology, 143,* pp.117-125.

- **Conference oral presentation:** "*Dose accumulation assessing the validity of reduced PTV margins in head-and-neck radiotherapy*" at the 2019 Engineering & Physical Sciences in Medicine Conference (EPSM), 29th October 2019, Perth, Australia. This is the premier annual event for medical physicists and biomedical engineers in Australia and New Zealand held annually

- **Accepted journal article:** Lowther, N.J., Marsh, S.H. and Louwe, R.J., 2020. Dose accumulation to assess the validity of treatment plans with reduced margins in radiotherapy of head and neck cancer. *In publication in Radiotherapy and Oncology*

- **Poster highlight presentation:** "*DIR based dose accumulation to validate reduced PTV and PRV margins in head-and-neck radiotherapy*" at the 2020 European Society for Radiotherapy & Oncology (ESTRO) conference, 4th April 2020, Vienna, Austria. This is the premier annual event for European professionals in the radiation oncology field

- **e-Poster:** "*Quantifying the dose accumulation uncertainty after DIR in head-and-neck radiotherapy*" at the 2020 European Society for Radiotherapy & Oncology (ESTRO) conference to be held from the 3rd – 7th April 2020 in Vienna, Austria. This is the premier annual event for European professionals in the radiation oncology field

SCHOLARSHIPS, GRANTS AND AWARDS

This research would not have been possible without the generous support and funding received from the following:

- **PhD research scholarship:** A Lottery Health Research grant, administered by the New Zealand Lottery Grants Board, acting by and through the Department of Internal Affairs

- **PhD research supplemental scholarship:** The Edward and Isabel Kidson Scholarship, administered by Universities New Zealand

- **Professional development grant:** A travel grant, administered by the Cancer Society of New Zealand Wellington Division, to attend the 2017 EPSM conference in Hobart, Australia

- **Professional development grant:** A travel grant, administered by the Canterbury Branch, Royal Society of New Zealand, to attend the 2018 EPSM conference in Adelaide, Australia

- **Professional development grant:** A travel grant, administered by Research for Life, Wellington Medical Research Foundation, to attend the 2018 EPSM conference in Adelaide, Australia

- **Professional development grant:** A Claude McCarthy Fellowship, administered by Universities New Zealand, to attend the 2019 ESTRO conference in Milan, Italy

- **Scientific award:** A travel grant, administered by the Australian Institute of Nuclear Science and Engineering (AINSE) society awarded to the best student abstract from the NZ branch of Australasian College of Physical Scientists & Engineers in Medicine (ACPSEM) at the 2018 EPSM conference in Adelaide, Australia

- **Scientific award:** A travel grant, administered by the Australasian College of Physical Scientists & Engineers in Medicine (ACPSEM) NZ branch awarded to two medical physics students presenting at the 2019 NZPEM conference in Auckland, New Zealand

- **Post Graduate Publishing Bursary grant:** A publication publishing grant, administered by the Canterbury Medical Research Foundation (CMRF) to prepare and submit papers for publication in journals of international standing

- **Professional development grant:** A travel grant, administered by the Cancer Society of New Zealand Wellington Division, to attend the 2019 EPSM conference in Perth, Australia

CONTENTS

FIGURES

TABLES

NOMENCLATURE

3DCRT	Three-dimensional conformal radiotherapy
AAPM	American Association of Physicists in Medicine
ART	Adaptive radiotherapy
AI	Artificial intelligence
BED	Biological equivalent dose
BS	Brainstem
B-spline	Basis spline
CBCT	Cone beam computed tomography
CI	Confidence interval
CT	Computed tomography
CTV	Clinical target volume
dCT	Deformed planning computed tomography
DFE	Displacement field error
DIR	Deformable image registration
DVF	Deformation vector field
DVH	Dose-volume histogram
EWMA	Exponentially weighted moving average
FOV	Field of view
GTV	Gross tumour volume
HD	High-dose
HN	Head-and-neck
HNRT	Head-and-neck radiotherapy
HPV	Human papillomavirus
HU	Hounsfield unit
IC	Inverse consistency
ICE	Inverse consistency error
ICRU	International Commission on Radiation Units and Measurements
ID	Intermediate-dose
IGRT	Image-guided radiotherapy
IMRT	Intensity modulated radiation therapy

LD	Low-dose
LR	Loco-regional
LRR	Loco-regional recurrence
MDT	Multidisciplinary team
MLC	Multi-leaf collimator
MRI	Magnetic resonance imaging
MU	Monitor unit
NTCP	Normal tissue complication probability
OAR	Organ at risk
pCT	Planning computed tomography
PG	Parotid gland
PRV	Planning risk volume
PTV	Planning target volume
ROI	Region of interest
SA	SmartAdapt
SC	Spinal cord
SCC	Squamous cell carcinoma
SMG	Submandibular gland
SPC	Statistical process control
TCP	Tumour control probability
TPS	Treatment planning system
VMAT	Volumetric modulated arc therapy
WBCC	Wellington Blood & Cancer Centre
WLS	Weighted least squares

1 INTRODUCTION

1.1 Overview

Emerging literature has reported reduced treatment-related toxicity after reducing treatment margins in head-and-neck radiotherapy (HNRT) [1–4]. Treatment margins are applied volumetrically around the tumour to ensure adequate coverage by accounting for various treatment related uncertainties. As it was considered whether it was possible to implement reduced treatment margins at the Wellington Blood & Cancer Centre (WBCC) with the aim to improve patients' treatment-related toxicity, it was recognised that many department-specific aspects of HNRT can influence treatment outcomes [5] and should all be considered when the treatment margins are amended.

The goal of the research detailed in this dissertation was to facilitate a safe treatment margin reduction in HNRT at the WBCC by developing a quality management "safety net" for patients who may be at risk of unacceptable loss in target coverage and/or overdosage to normal tissues with reduced treatment margins. The quality management "safety net" developed in this research was underpinned by deformable image registration (DIR). Background information regarding HNRT, DIR and current practices at the WBCC is provided in this *Introduction* chapter to aid comprehension of the three linked studies which constitute this dissertation.

1.2 Head-and-neck cancer

Head-and-neck (HN) cancers are a group of malignancies that include tumours of the nose and sinuses, orbits, salivary glands, thyroid and parathyroid glands, skin, lymph nodes and mucous membranes. Mucous membranes line structures in the HN region such as the nasopharynx, oropharynx, oral cavity, hypopharynx and larynx (Figure 1.1). Squamous cell carcinoma (SCC) of the mucous membranes

comprise approximately 95% of HN cancers [6–8]. The most important risk factors are tobacco, alcohol and more recently, the human papillomavirus (HPV) infection [7,9–11]. At presentation, about two-thirds of patients with HN SCC demonstrate advanced stage disease although incidence of metastasis is infrequent [12]. Patients that survive loco-regional (LR) advanced disease face an increased risk of mortality from cardiac and respiratory illnesses [6,13–15]. In addition, these patients also face an increased risk of second primary tumours which are often related to tobacco [6,13–15]. In 2016, New Zealand recorded 520 new HN cancers of the lip, oral cavity and pharynx with incidences in men constituting over 65% of these cases [16]. Despite HN cancers accounting for only about 2% of newly registered cancer cases [16], a five-year relative survival rate of approximately 63% [17] indicates a significant health issue which is complex to treat. In comparison, the five-year relative survival rate of prostate cancer and female breast cancer is approximately 91% and 87%, respectively [17]. For brain and pancreatic cancer, this rate drops to approximately 19% and 5%, respectively [17]. Radiotherapy as either the sole treatment modality, concurrent with chemotherapy or adjuvant following surgery is an extremely effectual treatment for HN cancers [18].

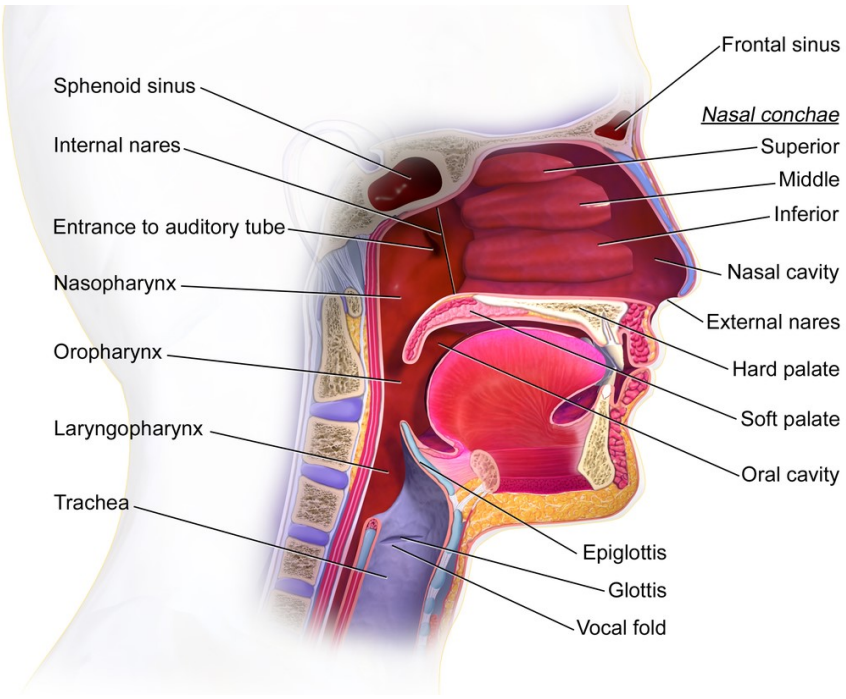


Figure 1.1: Structures of the head-and-neck (HN) region. Reused under a Creative Commons licence, Blausen.com staff (2014). "Medical gallery of Blausen Medical 2014". WikiJournal of Medicine 1 (2). DOI:10.15347/wjm/2014.010

1.3 Radiotherapy for head-and-neck cancers

Radiotherapy utilises a linear accelerator (Figure 1.2) to deliver radiation dose to a patient. Radiotherapy aims to deliver high radiation doses to tumour volumes and/or areas of suspected microscopic malignant disease whilst attempting to spare adjacent healthy tissues. Radiation is delivered to target volumes through surrounding normal tissues and as a result, the normal tissues cannot be completely spared.

Figure 1.2: A linear accelerator. Radiation is emitted through the head of the linear accelerator and deposited in the patient. Revised and reproduced with permission: for the National Cancer Institute © (2016) Terese Winslow LLC, U.S. Govt. has certain rights.

Tumour control probability (TCP) and normal tissue complication probability (NTCP) models are often used to describe the probability of achieving local control and corresponding normal tissue toxicity for a given radiation dose [19]. Maximising the TCP while minimising the NTCP is the optimal radiation delivery regime. The therapeutic ratio generally refers to the ratio of TCP to NTCP at an arbitrary level of response (typically 0.05) for normal tissue (Figure 1.3). Typically, an appropriate radiotherapy treatment has a TCP ≥ 0.5 and an NTCP ≤ 0.05 [19]. A higher therapeutic ratio is present when the NTCP curve is further to the right of the TCP (Figure 1.3). A higher therapeutic ratio [20] makes it easier to achieve the radiotherapeutic goal and reduce the likelihood of treatment complications. Radiotherapy dose distributions that can conform tightly to target volumes to spare healthy tissues as much as possible will maximise the therapeutic ratio.

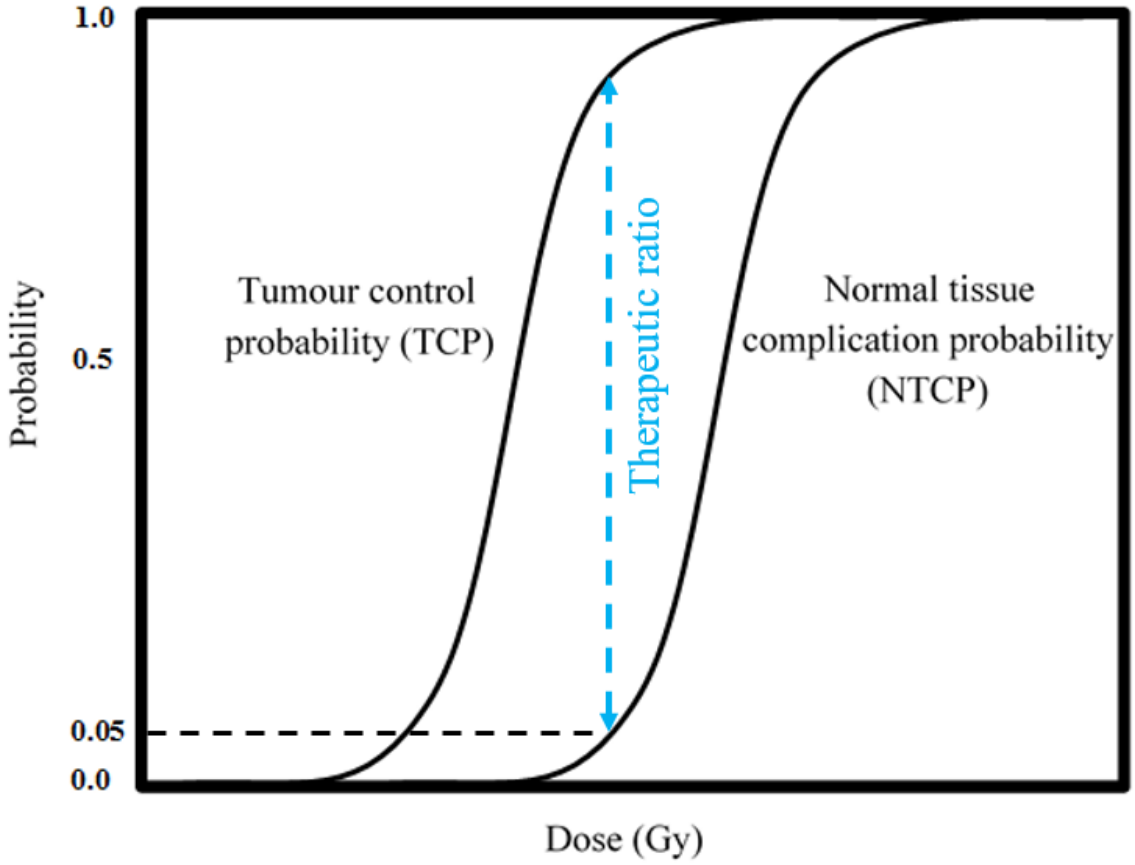


Figure 1.3: The principle of the therapeutic ratio. The therapeutic ratio is the ratio of tumour control probability (TCP) to normal tissue complication probability (NTCP) at an arbitrary probability level for normal tissue (typically 0.05).

The introduction of intensity modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT) in the last decade has enabled dose deliveries that conform more tightly to tumour volumes compared to the previously applied treatments using three-dimensional conformal radiotherapy (3DCRT) [21–23]. The principle of 3DCRT involves shaping a dose distribution to the target as prescribed while avoiding normal tissues as much as can be achieved. For 3DCRT, this is most commonly accomplished through the application of a limited number of radiation beams (typically 3 – 9 [24,25]) at static gantry angles, the use of lead inserts or multi-leaf collimators (MLC), and through treatment planning dose calculations. The treatment planning dose calculations are based on radiation transport models [26] which are applied to 3D anatomical patient data from computed tomography (CT) imaging. The treatment planning dose calculations are performed using a treatment planning system (TPS) [19]. A TPS is computer software that is used in external beam radiotherapy to generate beam

shapes and dose distributions with intent to maximise tumour control and minimise normal tissue complications [19]. MLCs are beam shaping devices made up of individual "leaves" that move independently from one another to shield normal tissues. The leaves are constructed from a high atomic number material, typically tungsten. IMRT is an approach more advanced than 3DCRT and is able to provide improved dose distribution conformity to irregularly shaped target volumes compared to 3DCRT, and has the ability to produce concave dose distributions [18]. Figure 1.4 illustrates the difference in dose distribution conformity between 3DCRT and IMRT. The increased conformity of IMRT is achieved by varying the photon fluence [27] across the beam aperture through using a large number of beam segments with a distinct shape for each static gantry angle. The progress in technological development of MLCs and inverse planning algorithms to optimise individual patient treatment plans was instrumental for the development of IMRT. VMAT also produces photon fluence variation across the beam aperture. However instead of delivering multiple beams at static gantry angles, VMAT delivers subsequent segments in a dynamic fashion by rotating the linear accelerator gantry while the gantry rotation speed, dose rate and MLC leaf movement speed are varied. The advantages of VMAT in head-and-neck radiotherapy (HNRT) are shorter delivery times, superior target dose homogeneity (double arc VMAT) and lower number of total monitor units (MUs), in comparison to IMRT [28,29]. Shorter delivery times imply a reduction in the risk of patient movement during treatment, and possibly a reduction in psychosocial factors attributed to being in the treatment room [30]. A reduced number of MUs (a measure of dose output from a clinical accelerator) indicates that equivalent treatment goals can be achieved with the overall production of less dose.

During the treatment preparation phase, the planning computed tomography (pCT) scan is acquired and used to delineate the following tumour target volumes as defined by the 1993 *International Commission on Radiation Units and Measurements (ICRU) Report 50 - Prescribing, Recording, and Reporting Photon Beam Therapy* [31]:

- The Gross Tumour Volume (GTV), the gross palpable or visible/demonstrable extent and location of malignant growth [31]

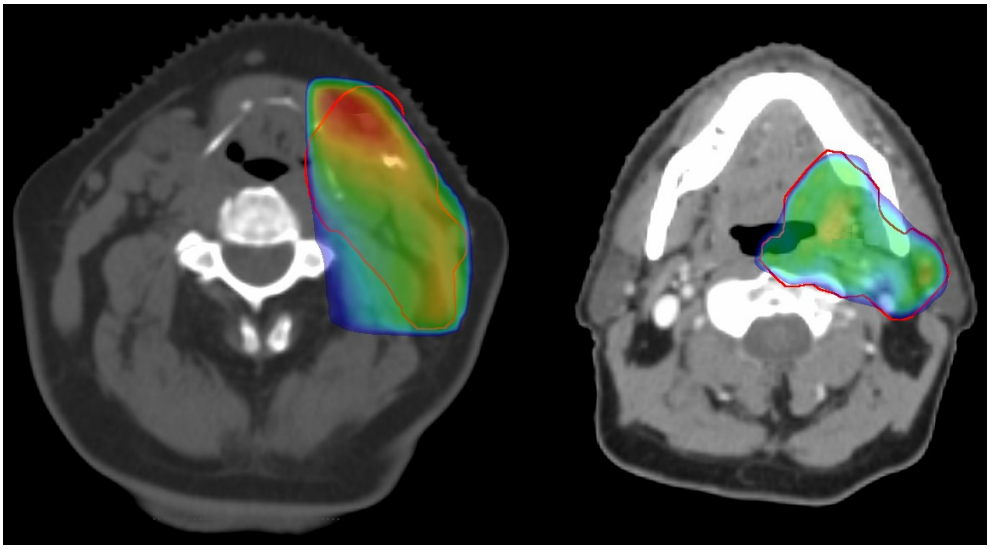


Figure 1.4: Colourwash representation of dose distributions' conformity to the target volumes (solid red contours). A three-dimensional conformal radiotherapy (3DCRT) plan (left) and a Volumetric Modulated Arc Therapy (VMAT) plan (right). The colourwash represents 95% of the prescribed dose. The 95% isodose conforms highly but not perfectly to the target volume for the VMAT plan in comparison to the conformity of the 3DCRT plan. Two different patients are shown.

- The Clinical Target Volume (CTV), a tissue volume that contains a demonstrable GTV and/or subclinical microscopic malignant disease, which has to be sterilised. This volume has to be irradiated to an adequate dose to achieve a successful treatment [31]. For HN cancers, two different approaches to delineating the primary tumour CTV exist, one based on geometric expansion of the GTV and the other one based on anatomical expansion of the GTV using compartmentalisation of HN anatomy [32–34]
- A Planning Target Volume (PTV), which consists of the CTV plus an expanded margin. The PTV margin ensures CTV coverage by accounting for various uncertainties including patient positioning, geometrical accuracy of the treatment machine and geometrical uncertainties of target volume definition. Most of these uncertainties can be quantified and applied in a so-called PTV margin recipe to estimate the appropriate PTV margin [35,36]. However, for HNRT it is difficult to quantify the uncertainty in target volume definition because a gold standard is seldom available. Studies directly comparing histopathological analysis with delineation of the GTV using pre-treatment imaging indicated that target delineation was the main source of uncertainty in the observed differences for laryngeal and hypopharyngeal cancers [37,38]. In

addition, it was noted that the inter- and intra-observer variation in clinician target delineation was considerable [37,38]. An example of inter-observer delineation variation at the WBCC is shown in Figure 1.5.

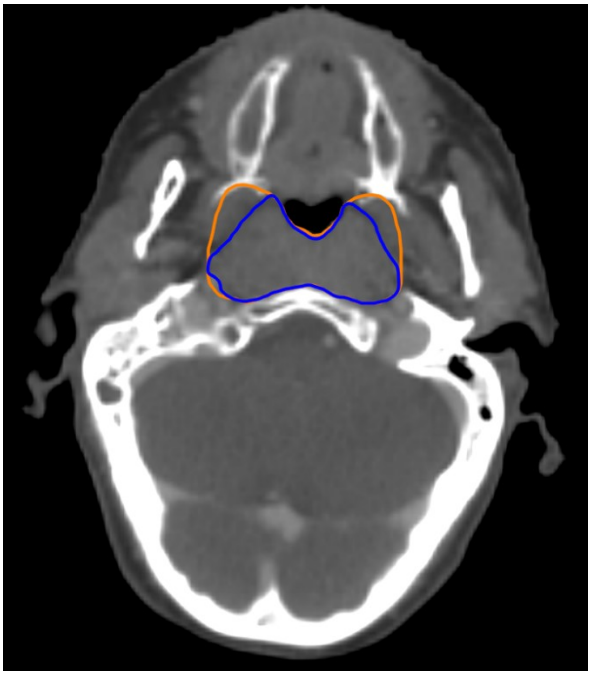


Figure 1.5: Blue and orange contours illustrating an example of inter-observer gross tumour volume (GTV) delineation variation of two clinicians at the Wellington Blood & Cancer Centre (WBCC).

A number of CTVs and PTVs may be delineated on a patient's pCT scan and these are irradiated to different dose levels depending on their risk [39–41]. Figure 1.6 shows the GTV relative to the high-dose CTV, the high-dose PTV, the low-dose CTV and the low-dose PTV. In addition to the target volumes, a number of patient- and treatment-specific organs at risk (OARs) are delineated. An OAR is defined as a normal tissue organ whose sensitivity and tolerance for radiation may result in unacceptable radiation treatment toxicities [19]. Typical OARs in HNRT include the brainstem (BS), spinal cord (SC), parotid glands (PGs), submandibular glands (SMGs), oral cavity, optic chiasm, optic nerves, pharyngeal constrictor muscles, laryngeal cavity and eyes. Similar to the PTV, a planning risk volume (PRV) takes the net effect of geometrical variations, uncertainties in OAR delineation and patient positioning into consideration [42,43]. The PRV is a geometrical concept, and is frequently used in the treatment plan optimisation for the critical serial OARs (i.e., the BS and SC) as opposed to the BS and

SC contours themselves. A treatment plan (beam arrangement, monitor units, fluence maps) is generated on the pCT and optimised according to target coverage [44], critical OAR avoidance and non-critical OAR sparing requirements [45].

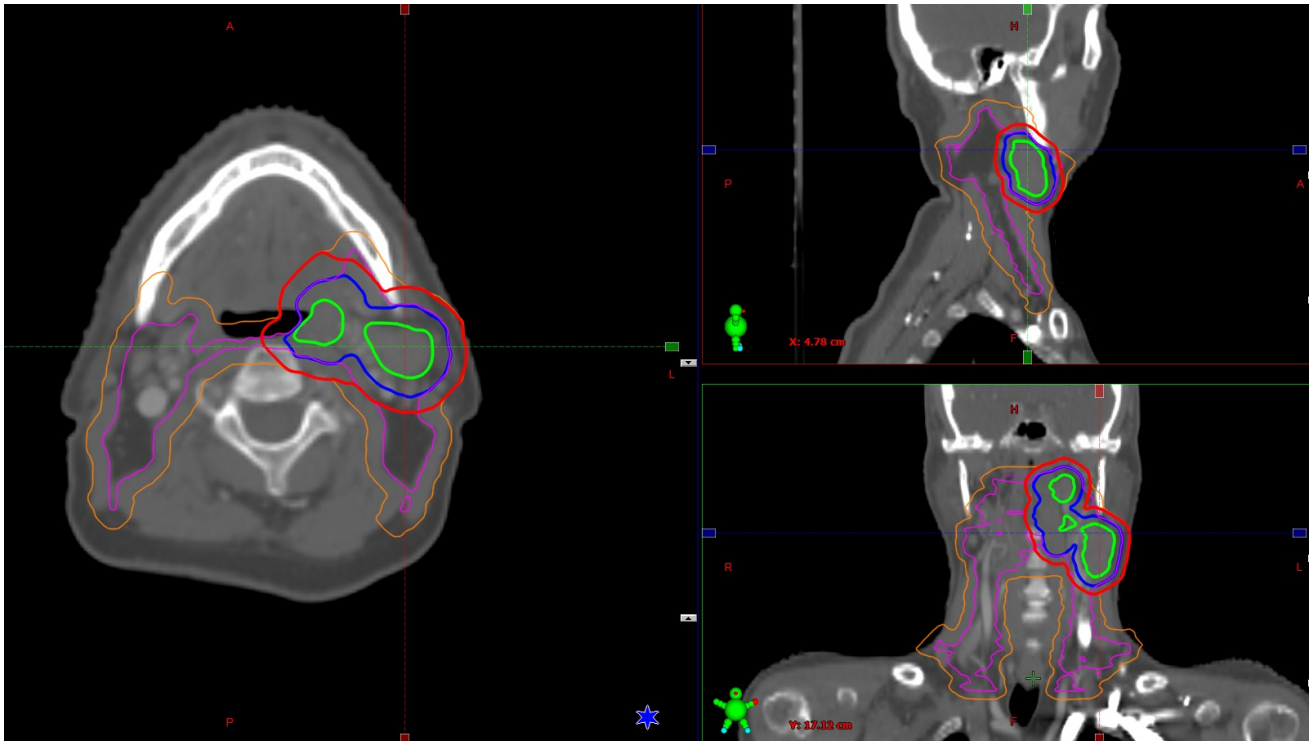


Figure 1.6: Target volumes shown relative to each other in the axial (left), sagittal (top right) and coronal (bottom right) planes. The gross tumour volume (GTV) is indicated by the bright green contour. The high-dose clinical target volume (CTV) in blue. The high-dose planning target volume (PTV) in red. The low-dose CTV and PTV are shown in purple and orange, respectively.

1.3.1 Dose-volume histograms

Dose-volume histograms (DVHs) [46] are an effective tool often used in radiotherapy to quantify the treatment plan's target coverage, critical OAR avoidance and non-critical OAR sparing [19]. A cumulative DVH is a frequency plot showing the volume receiving a dose greater than or equal to a given dose, plotted against dose. DVHs are typically displayed in the form of percentage of total volume [46]. Example DVHs are shown in Figure 1.7. Structure-specific DVH quality metrics are used to ensure that the treatment plan provides appropriate target coverage, critical OAR avoidance and non-critical OAR sparing. For PTVs, the $D_{98\%}$ achieving at least 95% of the prescribed dose is commonly used as the DVH quality metric [29,47] during plan optimisation and evaluation. That is, at least 98% of the

PTV should receive at least 95% of the prescribed dose. This requirement is one of the many requisites of an acceptable treatment plan. The orange filled circle in Figure 1.7 shows the high-dose PTV's planned $D_{98\%}$ which is equal to 63.8 Gy. Considering that the prescribed dose was 66 Gy, the planned $D_{98\%}$ for the high-dose PTV must exceed 62.7 Gy. The DVH quality metric typically used for critical OARs (i.e., SC and BS PRVs) is the max dose D_{max} (Figure 1.7) or the minimum dose to 0.1 cm^3 of the volume that receives the highest dose, $D_{0.1cc}$ [45]. For non-critical OARs, the DVH quality metric is typically the mean dose received by the structure, D_{mean} [45].

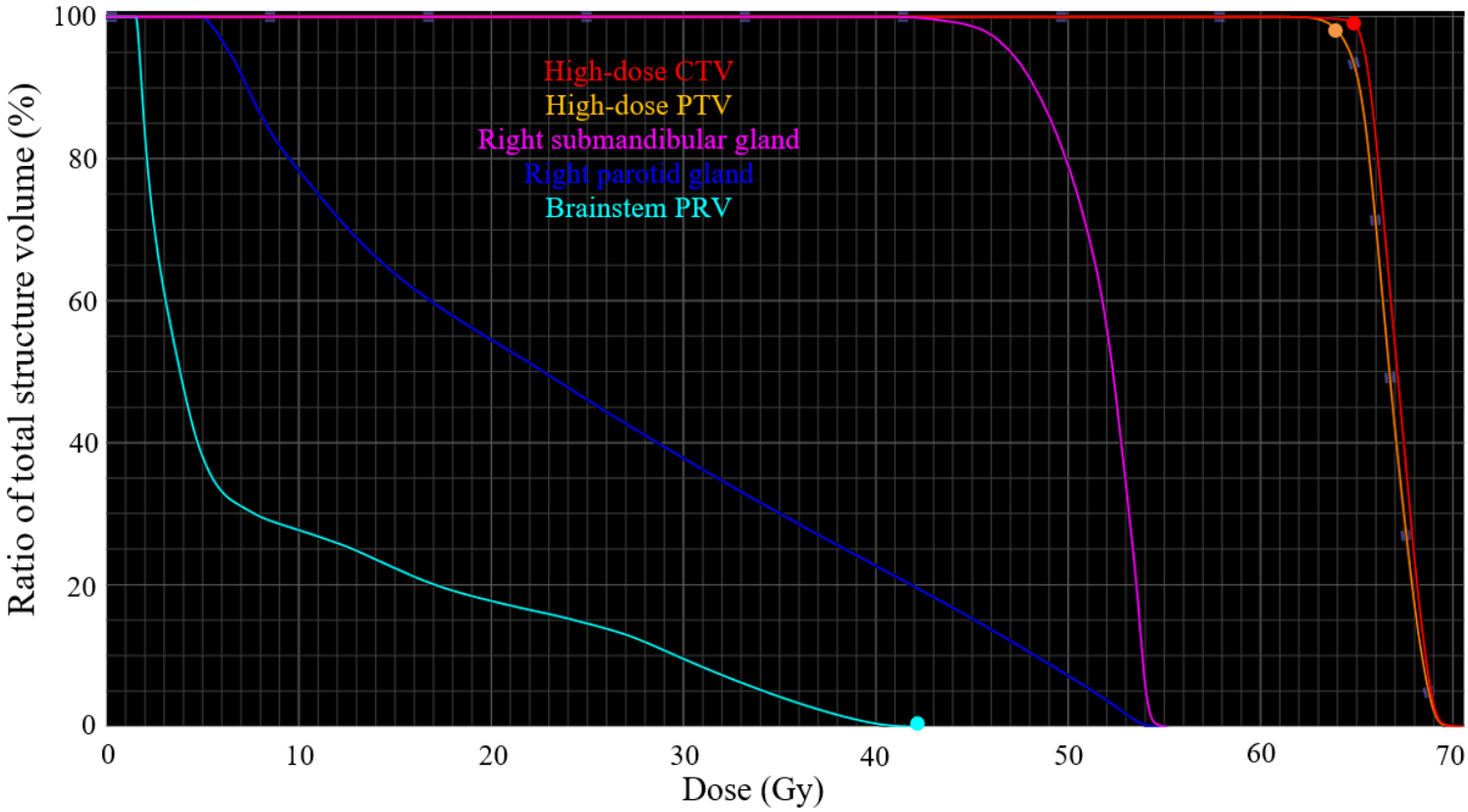


Figure 1.7: Selected dose-volume histograms (DVHs) for an example patient's planned dose. The red filled circle is the high-dose (HD) clinical target volume (CTV) $D_{99\%}$. The orange filled circle is the HD planning target volume (PTV) $D_{98\%}$. The teal filled circle is the brainstem planning risk volume (PRV) D_{max}. $D_{99\%}$ = dose to 99% of the volume. $D_{98\%}$ = dose to 98% of the volume. D_{max} = max dose.

1.3.2 Radiotherapy toxicity

In general, radiation toxicity can be separated into two groups: stochastic and deterministic. With increasing dose, stochastic effects increase in occurrence probability but do not increase in severity

[19]. An example of a stochastic effect is the induction of a secondary cancer [48]. Even at low doses there is a probability of a stochastic effect occurrence and as such, stochastic effects do not have a threshold dose below which they do not occur [19]. Deterministic effects increase in severity with increasing dose, often above a threshold dose. The remainder of *Sub-section* 1.3.2 details deterministic toxicities.

One of the main challenges in HNRT is reducing toxicity for the patient while maintaining tumour control due to the intimately located target volumes and normal tissues. Radiation-induced toxicity for HN cancer patients has a deleterious and well-documented impact on a patient's quality of life following treatment [49–51]. Despite significant improvements in a patient's quality of life after the introduction of modulated radiotherapy techniques compared to 3DCRT [52,53], continued efforts should be made to further improve quality of life for all patients during and after HNRT. Toxicities as a result of HNRT are separated into early and late adverse effects. Early effects are characterised as occurring during or immediately after HNRT and mostly decrease in severity over time. Conversely, late effects are generally considered irreversible and persist for months or years after completion of HNRT [49]. It should be realised that the recent introduction of IMRT and VMAT implies that studies reporting on the late effect toxicities associated with these techniques are limited.

Dryness of mouth, oral pain and difficulties with speech are characteristic symptoms of the most common HNRT late toxicity, xerostomia. Xerostomia is induced as a result of salivary gland (PG, SMG, sublingual and minor) irradiation during HNRT [54–56]. Figure 1.8 visualises the major salivary glands in the HN region. It is well-documented and well-known that a reduction in xerostomia can be achieved through PG sparing [25,55,57–62]. In contrast to the PGs, studies reporting on xerostomia endpoints and dose reduction strategies pertaining to the SMGs, sublingual and minor salivary glands are scarce.

Dysphagia is defined as difficulty in swallowing and commonly presents in the oral cavity, oropharynx and oesophagus [63,64]. Dysphagia can present as an early adverse effect through soft tissue inflammation, oedema, mucous production and xerostomia. Permanent reduction of swallowing

function even after healing of the early adverse effects can occur due to soft tissue fibrosis, scar tissue formation and neurological impairment [65]. A reduction in dose to the swallowing structures (pharyngeal constrictor muscles, glottis, supraglottis, and oesophagus) has been directly linked to a reduction in patient-experienced dysphagia [66–68].

Osteoradionecrosis [69], particularly of the mandible, is another documented adverse effect of HNRT and can lead to pain, loss of bone mineral and skin fistulas [70,71]. Hearing loss as a result of treatment is less-documented, however a small number of studies seeking to determine the underlying mechanisms of radiation-induced otological toxicity recommended consideration of dose to the inner ear [72,73]. Finally, a reduced opening of the jaw, known as trismus, is another recognised adverse effect of HNRT [74].

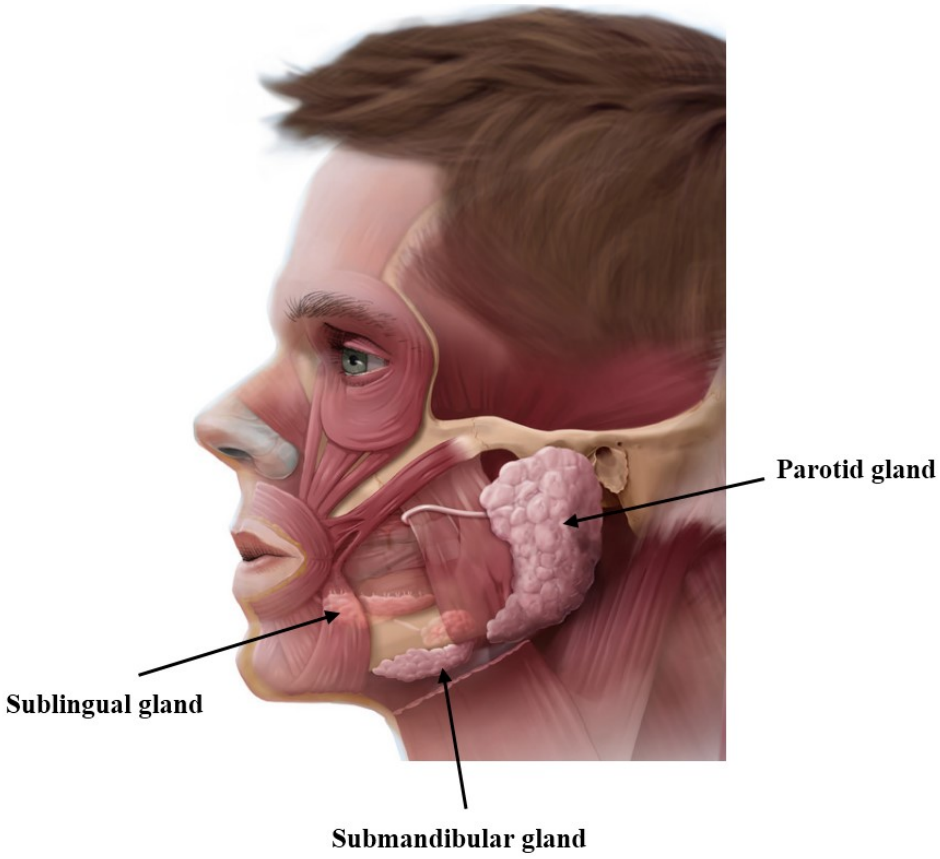


Figure 1.8: Major salivary glands of the head and neck. Modified under a Creative Commons licence, from the University of Dundee.

1.4 Anatomical changes

For HNRT, the length of treatment typically consists of daily treatment fractions from Monday to Friday for 4-7 weeks. The fractionation regime is based on radiobiological models of cancer and normal tissues. Fractionation takes advantage of the increased survival ability of normal tissues compared to cancer cells considering that normal cells are able to more effectively repair radiation damage if the total dose is split into smaller doses [75–77]. Normal tissue cells proliferate at a slower rate than cancer cells consequently having more time to repair damage before replication [75–77]. During each treatment fraction, the position of the patient must be carefully reproduced and patient-specific anatomy must be carefully monitored due to the intimately located target volumes and normal tissues. The dynamic nature of changes within the human body during the treatment period implies that the pCT acquired during the treatment preparation phase may not accurately represent the patient's anatomy during treatment [78]. Rigid setup errors and non-rigid deformations (which include non-rigid setup errors) of a patient's anatomy relative to the pCT anatomy have the potential to adversely affect the planned dose distribution [79,80].

1.4.1 Setup errors

The most common method of reducing patient positioning setup errors [81–85] in HNRT is to immobilise the patient using a head support and a thermoplastic fixation mask, and apply treatment couch corrections to correct any rigid patient setup deviations. Generally, treatment couch corrections are carried out immediately prior to delivery of each treatment fraction, based on the registration of an in-room imaging modality scan to the pCT. This technique of imaging a patient immediately prior to delivery of their treatment fraction and correcting for rigid patient setup errors is known as image guided radiotherapy (IGRT) [86]. Typically, the in-room imaging modality is a cone beam CT (CBCT). CBCT integrates CT imaging with a linear accelerator and involves acquiring multiple planar images from a kilovoltage or megavoltage cone beam of photons rotating about the patient on the treatment couch [19]. Residual setup errors that cannot be corrected for with rigid couch movements are known as non-rigid deformations. Figure 1.9 illustrates the correction of setup errors.

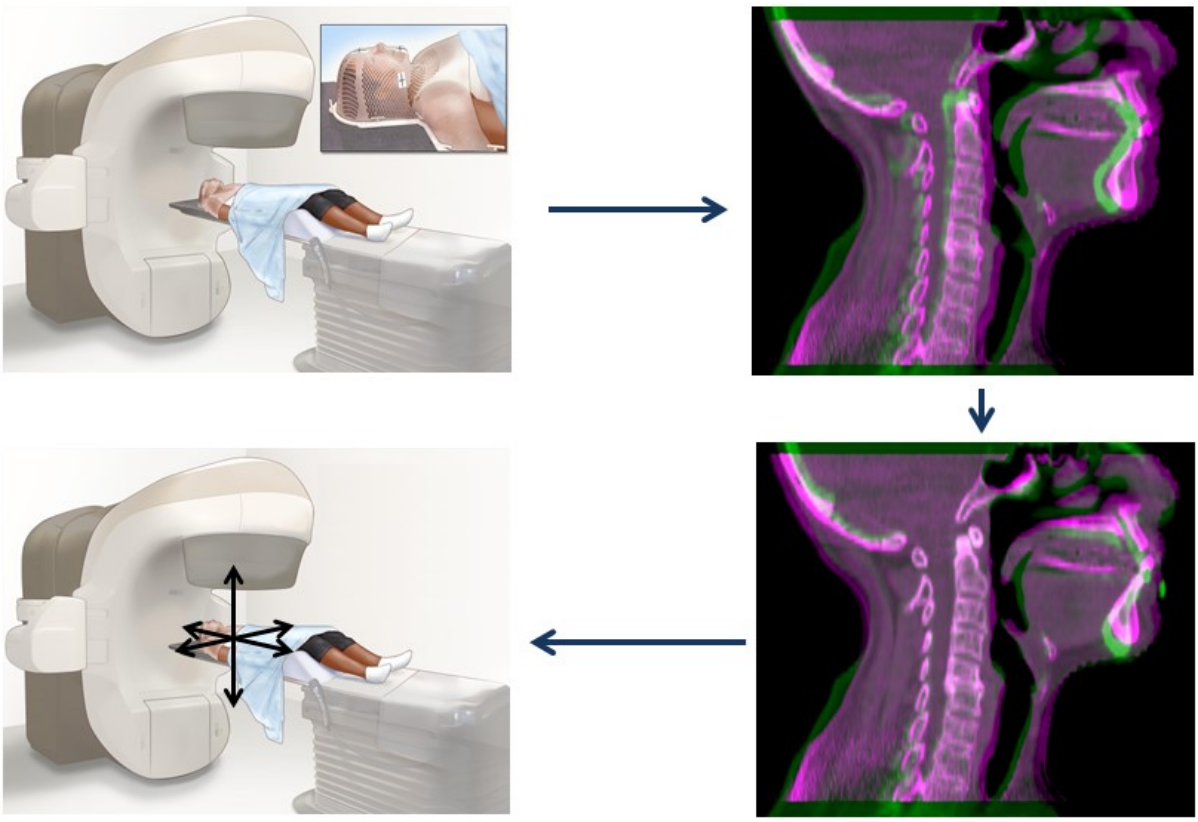


Figure 1.9: Reduction of setup errors in image guided radiotherapy (IGRT): (top left) The patient is positioned on the treatment couch with a head support and thermoplastic fixation mask in an attempt to reproduce the planning computed tomography (pCT) scan position; (top right) A cone beam CT (CBCT) scan is taken (magenta) and digitally overlaid on the pCT (green); (bottom right) The CBCT and pCT are digitally matched by treatment staff prioritising a region of interest. In this case, the C1-C3 vertebral bodies; (bottom left) the digital match is applied to the treatment couch in three translational degrees of freedom. Note that an additional three rotational degrees of freedom can be used. Residual setup errors (bottom right) seen at the mandible and hyoid are non-rigid deformations. Top and bottom left panes revised and reproduced with permission: for the National Cancer Institute © (2016) Terese Winslow LLC, U.S. Govt. has certain rights.

1.4.2 Non-rigid deformations

Non-rigid deformations include non-rigid patient positioning change, weight change, OAR shift and OAR shrinkage. Examples of non-rigid deformations that can occur during HNRT are illustrated in Figure 1.10 by means of a patient's pCT overlaid with an on-treatment CBCT.

Table 1 summarises non-rigid patient positioning as reported in the literature and reports the group mean, *M,* in accordance with population statistics described in the Appendix of Remeijer et al. [87]. The largest non-rigid patient positioning deformations (i.e., *M*) were typically observed for the

mandible. Non-rigid patient positioning deformations were calculated relative to C1-C3 or a similar structure. The selection of this reference structure provided a clinically relevant measure of deformation considering its common prioritisation in HNRT (Figure 1.9). A comparison of the literature reported dispersion of the systematic deformation $\sum$ and random deformation σ can be found in a previous study conducted at the WBCC [88].

Weight change is commonly recorded during HNRT treatment to monitor the overall health status of the patient [89–92]. However, a patient's overall weight change is not necessarily an indicator of volume change in the HNRT treatment region considering that weight loss may not be evenly distributed over the entire body. A number of investigations have reported on the change in volume and neck diameter of various HN sub-regions [78,93–95]. These studies then correlated the HN sub-regions' contour change with total weight change. For example, Barker et al. [78] found a significant correlation (ρ = 0.92; $p < 0.01$) between patient weight and external volume (one CT slice) at the level of C2 in their fourteen patient study. However, these investigations [78,93–95] do not account for changes in patient positioning between planning and during-treatment images. Patient tissue could enter/exit the measured zone due to patient positioning inconsistencies and not represent true treatment region contour change. To further clarify, a tilt of the head not present at planning may cause tissue to leave the measured HN sub-volume (or neck diameter plane) when no total change of patient weight is observed, a false positive for weight loss. There is no true gold standard to quantify anatomical changes in the HN treatment region solely due to weight change.

Brouwer et al. [79] have comprehensively summarised 51 studies reporting on anatomic and dosimetric changes in HN OARs during radiotherapy. The majority of papers ($n = 38$) in their review included anatomical changes to the PG. The high number of investigations can be attributed to the PG's well-known dose relationship with xerostomia and saliva production [56,59]. For the 38 reviewed studies, the average volume reduction of the PGs was 26 ± 11% and PGs tended to migrate medially during treatment [79].

*Table 1: Group mean deformation, M, relative to C1-C3 or similar structure**

M_{ROI} (mm)	**C3-C5** 3D dev.				**C5-C7** 3D dev.				**C7-caudel** 3D dev.				**Mandible** 3D dev.				**Maxilla** 3D dev.				**Occipital** 3D dev.				**Larynx** 3D dev.			
	vrt	*lng*	*lat*	***3D***	*vrt*	*lng*	*lat*	***3D***	*vrt*	*lng*	*lat*	***3D***	*vrt*	*lng*	*lat*	***3D***	*vrt*	*lng*	*lat*	***3D***	*vrt*	*lng*	*lat*	***3D***	*vrt*	*lng*	*lat*	***3D***
WBCC [88]	-0.3	-0.1	0.0	**1.6**	-0.2	-0.1	-0.2	**2.3**	0.4	0.1	-0.3	**2.9**	1.0	-0.7	-0.2	**2.9**	-	-	-	-	-0.4	0.3	0.2	**2.2**	0.0	0.4	0.1	**3.2**
Polat et al. [82]	-	-	-	**3.0**	-	-	-	**3.0**	-	-	-	-	-	-	-	**3.2**	-	-	-	-	-	-	-	**2.7**	-	-	-	-
Van Kranen et al. [83]	0.1	0.1	0.1	-	0.2	-0.2	0.4	-	0.8	0.5	0.9	-	0.3	1.3	0.0	-	-	-	-	-	0.3	0.6	0.4	-	0.2	-0.3	0.5	-
Giske et al. [96]	-	-	-	-	1.2	0.3	0.1	-	2.0	0.3	0.0	-	0.3	-0.6	0.0	-	0.2	-0.6	0.0	-	-	-	-	-	1.7	-0.5	0.2	-
Graff et al. [81]	-	-	-	-	-	-	-	-	-	-	-	**2.7**	-	-	-	**4.0**	-	-	-	**3.5**	-	-	-	-	-	-	-	-
Zhang et al. [84]	-0.5	0.9	-0.7		-0.3	1.1	-1.0	-	-	-	-	-	-	-	-	-	0.4	1.9	-1.2	-	-	-	-	-	-	-	-	-

*The regions of interest (ROIs) used in WBCC [88] and van Kranen et al. [83] are the same however, other studies used varying sub regions which were compared to the ROIs of Table 1's column header. Polat et al. [82] matched on C4-C6 and the whole skull, which were compared to the results for C3-C5/C5-C7 and the occipital bone, respectively. Giske et al. [96] and Graff et al. [81] both used C1-C2 as the reference structure for the calculation of deformation as opposed to C1-C3. Zhang et al. [84] used the marked isocentre on the immobilisation mask as the reference structure. The results for Giske et al. [96] presented for C6, T2 and left maxillary sinus were compared to C5-C7, C7-caudal and maxilla, respectively. Graff et al. [81] included C7-T1, which was also compared to C7-caudal. The results for Zhang et al. [84] presented for C2, C6 and palatine process of the maxilla were compared to C3-C5, C5-C7 and maxilla, respectively.

Only three studies in the Brouwer et al. [79] review investigated the SMGs which is in line with the general scarcity of SMG investigations in the context of HNRT. On average, a SMG volume reduction of 22% (range; 15 – 32%) during HNRT was reported.

In general, the volume of the brainstem and spinal cord remained constant during HNRT [97]. Swelling of the larynx and/or pharyngeal constrictor muscles was found in two studies after at least five weeks of treatment [98,99]. An increase in the width of the pharyngeal constrictors to 111% of their width at planning was observed in Popovtzer et al. [100] for those structures receiving more than 50 Gy.

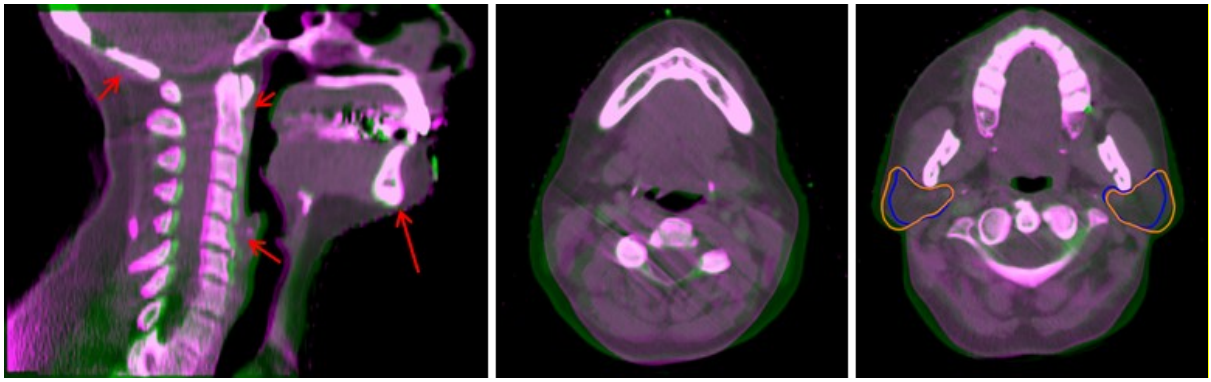


Figure 1.10: Examples of non-rigid deformations that can occur during head-and-neck radiotherapy (HNRT). The planning computed tomography (pCT) scan (green) overlaid with a during-treatment cone beam computed tomography (CBCT) scan (magenta). (Left) Non-rigid patient positioning characterised by differential movement of bony regions. (Middle) Weight loss resulting in a volume reduction of the treatment region, shown by a shrinkage of the external contour in the posterior region. (Right) Parotid gland shrinkage and shift shown by the pCT (orange) and CBCT (blue) contours.

1.4.3 Dosimetric consequences

IMRT has reported improved quality of life, increased sparing of salivary glands and no change in two- and three-year local-regional control rates for HNRT, compared to 3DCRT [52,53,61,62,101,102]. However considering the highly conformal dose distributions (Figure 1.4) of modulated radiotherapy (i.e., IMRT and VMAT), a small change in patient anatomy can lead to target volume underdosage and/or unacceptably high dose to normal tissues [58,93,103–110]. In the Brouwer et al. [79] review, 24 papers reported on the dosimetric consequences to the PGs as a results of anatomical changes during HNRT. An average PG D_{mean} increase of 2.2 ± 2.6 Gy relative to the planned D_{mean} was observed across

these studies. The largest average PG D_{mean} increase for a single study was 10.4 Gy from the study of Chen et al. [111]. Only one paper reported changes to the delivered SMG D_{mean} relative to the planned D_{mean}. An approximated actually delivered SMG D_{mean} increase of 0.9 Gy relative to the planned SMG D_{mean} (i.e., 52.8 vs. 51.9 Gy) was reported. Not included in the Brouwer et al. [79] review, Dijekma et al. [55] found a significant correlation between grade 4 or 5 night-time xerostomia and a reduction in SMG saliva production, caused by SMG irradiation. Changes to D_{max}, $D_{1\%}$ or $D_{2\%}$ of the spinal cord and/or brainstem was reported by 15 studies in Brouwer et al. [79] which observed an average increase of less than 1.5 Gy. It must be stated that the chronological end-points in these studies were largely determined by availability of re-CT images and were not always at the end of treatment. Zhao et al. [112] observed the largest increase in $D_{1\%}$ over treatment which was 5.6 and 2.5 Gy for the SC and the BS, respectively. The relatively large number of papers reviewed by Brouwer et al. [79] has started a consensus on the expected dosimetric impact to OARs in the presence of anatomical changes during HNRT. However, the notable inter-study heterogeneity restricted completely unambiguous conclusions. In addition, it must be realised that patient-specific local dose gradients, patient-specific OAR geometry relative to target volumes and additional department-specific treatment factors [5] will ultimately cause non-uniform deviations in dose between patients. This could further explain the variation in results reported by Brouwer et al. [79].

1.5 Adaptive radiotherapy

The planned dose distribution is frequently undermined in HNRT due to anatomical changes (*Sub-section* 1.4.2) occurring during treatment. Adaptive radiotherapy (ART) is the process of administering accurate and precise treatment in the presence of these anatomical changes [113–117]. As described by Yan et al. [116], a comprehensive ART implementation should include four key steps: *(1)* Treatment dose assessment, the use of on-board imaging in combination with the therapy machine output to provide feedback of the accumulated dose in structures of interest at a given time point during treatment; *(2)* Treatment variation identification/evaluation, the multiple-role step that characterises random variation in treatment processes, estimates the total treatment dose, estimates the biological equivalent

dose (BED) [118] in structures of interest and the corresponding potential dose response, and provides feedback information in parametric form for treatment modification decisions; *(3)* Treatment modification decisions, a set of premeditated control rules. The control rules are triggered by the outputs of step *(2)* and used to determine whether the treatment needs modification and what form of modification is required. Typical modifications may be a patient positioning correction, a beam aperture correction, a new PTV or a treatment plan modification; (4) Adaptive treatment modification, application of the treatment modification which may be performed online or offline [116].

Typically, ART in HNRT is based on an ad-hoc re-plan (not necessarily generated on a re-scan CT) when a patient experiences considerable anatomical change [93,119–121]. This form of ART is known as triggered adaption and can be considered the most common form of offline adaptation in radiotherapy [114]. In general, triggered adaption is often based on subjective criteria [93,119–121] such as inadequate fixation equipment or anatomical changes which are deemed clinically relevant on visual examination [114]. A number of alternative offline and online strategies have been proposed for ART such as an average anatomy model [122,123], a library of plans [124–126], a scheduled adaption [127] or online re-planning [128]. All of the ART methods mentioned so far have finite accuracy and as a result, residual uncertainties will still be present [114]. It is likely that the conceptually simple but technologically complex real-time ART [113] will assist the mitigation of these residual uncertainties in the future.

1.6 Deformable image registration

Effective and precise HNRT relies on the ability to incorporate patient information from multiple time-points during treatment [129]. Image registration is the procedure of aligning different intra- or even inter-patient datasets to a single frame of reference and is abundant in radiotherapy [130].

Image registration algorithms are underpinned by three components:

- The transform, applied to the *source* image to align it to the *target* image. Typically, the *source* image represents the original anatomy (e.g., the pCT) and the *target* image represents anatomy during treatment (e.g., a CBCT). The transform can be simple, representing a rigid body with six transform parameters (i.e., translation along the 3 cardinal axes – sagittal, coronal and axial – and rotation about these axes: pitch, yaw and roll, respectively) or complex, representing a higher dimensional registration method where the number of transform parameters can equal up to three times the number of voxels in the *source* image (i.e., a unique 3D transform for each voxel) [130]
- A registration metric, the quantification of the degree to which two imaging datasets are aligned. A registration metric is typically classified as either geometry-based or intensity-based. Geometry-based metrics make use of anatomical features from the image data such as bony landmarks or organ boundaries while intensity-based metrics use the voxel intensity data directly [130]. The registration metric is minimised or maximised where the minimisation or maximisation corresponds to the best alignment of the *source* and *target* image
- An optimiser, which seeks the optimal set of transformation parameters to obtain the best image alignment as quantified by the registration metric. Considering the large number of voxels in medical images, an efficient optimiser must intelligently search subsets of the parameter space while also searching the whole parameter space to find the most accurate solution [130]

When a patient has non-rigid deformation during treatment (*Sub-section* 1.4.2) and two image sets cannot be sufficiently registered with rigid translations and rotations only, there may be a need to utilise DIR (a higher dimension registration method). Approaches to DIR can broadly be characterised as geometric, such as the B-spline registration (*Sub-section* 1.6.1) or physical, such as the demons registration (*Sub-section* 1.6.2) [130]. The unique 3D transforms for each voxel are often smoothed and regularised to generate an overall transform which is considered to be biologically realistic. The overall transform in DIR which models the direction and magnitude of tissue progression is referred to as the deformation vector field (DVF) [131]. The deformation described by the DVF can be used to propagate

contours from the *source* image to the *target* image (*Sub-section* 1.6.3), and in dose accumulation (*Sub-section* 1.6.4).

1.6.1 B-spline registration

A B-spline DIR transformation generates the overall non-rigid deformation (i.e., the DVF) from a set of weighted basis functions [132]. Functions in a set are named basis functions if every continuous function in the function space can be represented as a linear combination of those functions in the set. A corresponding set of control points are marked on the *source* and *target* images to generate the basis functions. The deformation at a given point in the image space is calculated as the weighted sum of the basis functions. Manipulating the weighting of each basis function affects only a local segment of the overall non-rigid deformation considering that each basis function is defined for a discrete sub-volume of the imaging space only [132]. Figure 1.11 visualises the B-spline deformation concept for a one-dimensional case. Increasing the number of control points allows more complex local non-rigid deformations to be modelled (Figure 1.11).

1.6.2 Demons registration

A demons DIR transformation generates the DVF from displacements of the individual voxels, or a combined set of these [132–134]. The original demons algorithm (below) and extensions of the original algorithm use voxel intensities to derive the estimated displacement $\vec{v}$ between the *source* and *target* images, and are based on the optical flow equation [133]. For an individual voxel, let the intensity in the source image be *s* and the intensity in the target image *t*.

$$\vec{v} = \frac{(s-t)\vec{\nabla}t}{\left(\vec{\nabla}t\right)^2 + (s-t)^2}$$

Where $\vec{v} = (v_x, v_y, v_z)$ and describes the voxel-to-voxel displacement in the three anatomical planes, and where $\vec{\nabla}t$ is the gradient of the *target* image and represents the relationship between neighbouring

points. The demons equation is typically solved iteratively by incremental changes in the position of voxels in the *target* image so that the *source* image can be morphed voxel-by-voxel into a spatial alignment with the *target* image [134,135]. In each iteration, the calculations of $\vec{v}$ are followed by a regularisation of the DVF which smooths noise and promotes geometric plausibility of the deformed image.

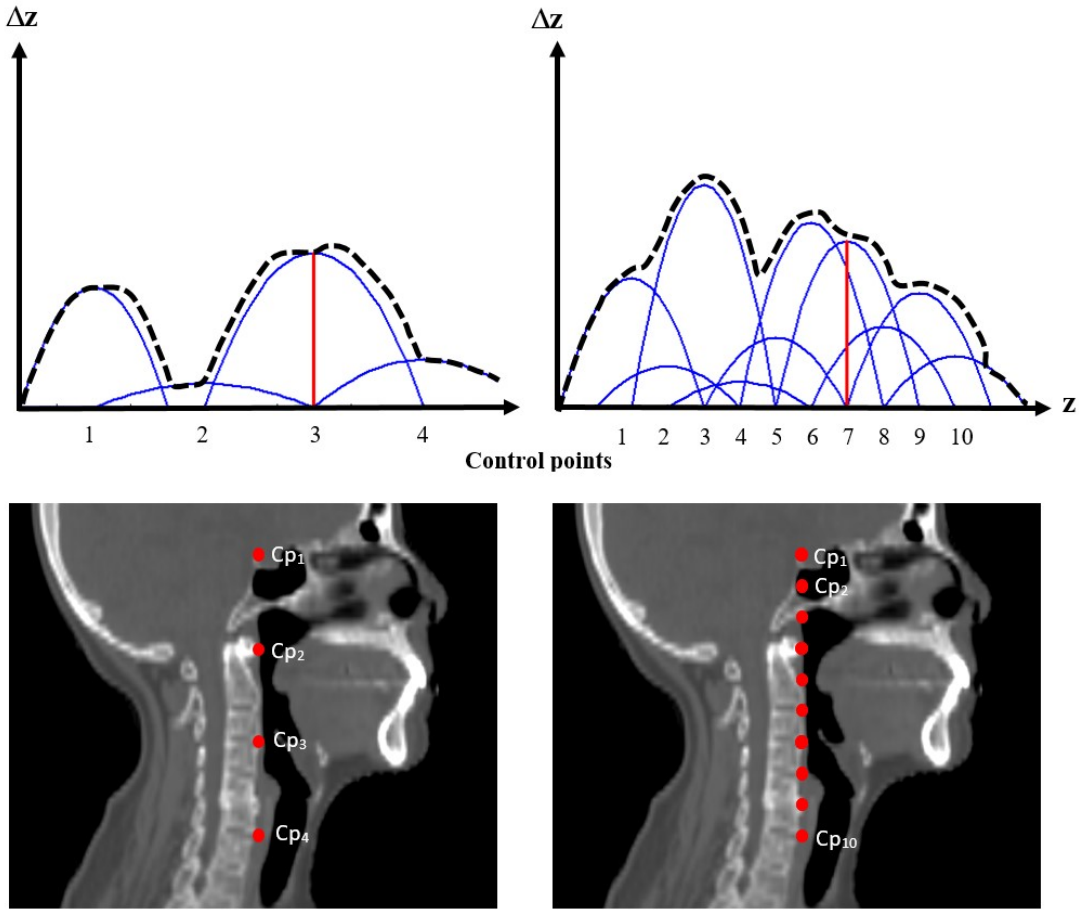


Figure 1.11: One-dimensional B-spline deformation concept in the superior-inferior (z direction). (Top plots) The overall non-rigid deformation (black dashed line) as a function of z is a weighted sum of the basis functions (blue lines) centred at a series of locations called control points (Cps). Increasing the density of Cps (right figures) aids modelling of complex local non-rigid deformations. Note the difference in overall non-rigid deformations between the top left and right plot. For the left and right plots, the weighting factor of the basis function (red vertical lines) affects the overall non-rigid deformation of two and four adjacent Cps, respectively.

1.6.3 Contour propagation

As described in *Section* 1.6, one rationale for aligning imaging data sets is to facilitate matching of information from one time point to another. Considering that the DVF models the direction and magnitude of tissue progression from the *source* to *target* image, contours of the *source* image can be deformed according to the DVF and copied to the target image. The process of deforming contours according to the DVF with subsequent copying to the *target* image is known as contour propagation [132,136–138]. Figure 1.12 illustrates the contour propagation procedure of pCT contours to a CBCT taken during treatment.

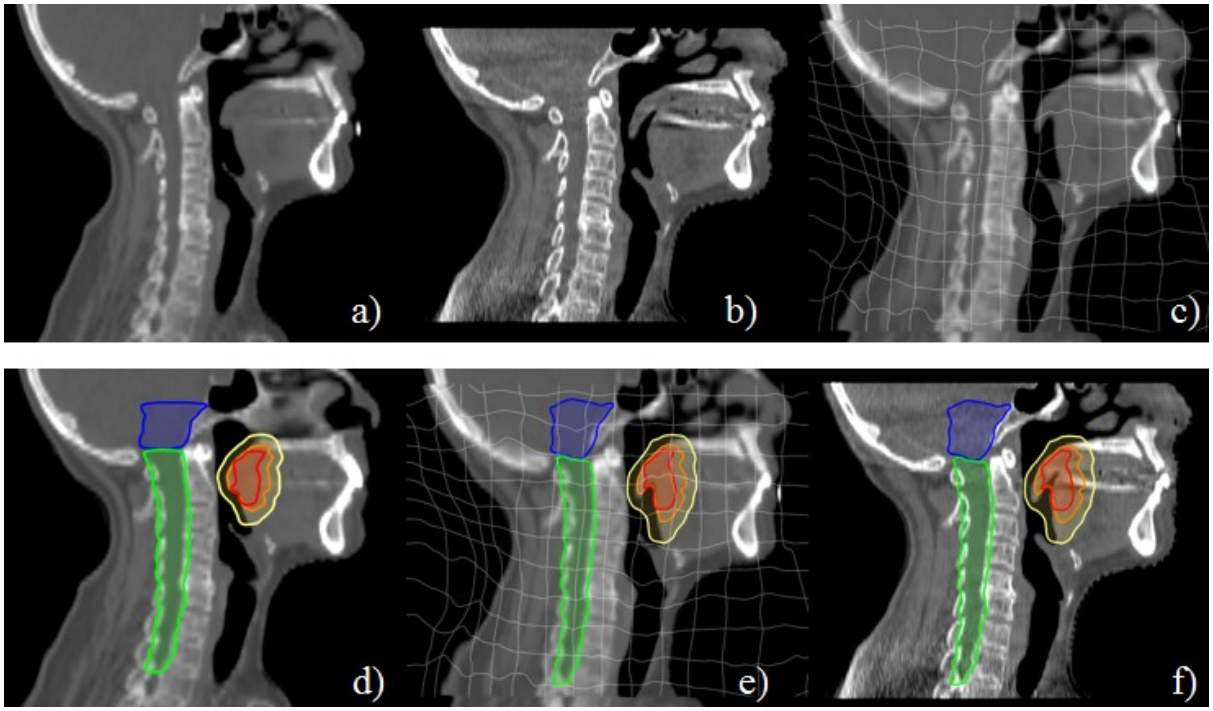


Figure 1.12: The contour propagation procedure. A planning computed tomography (pCT) scan (a) is matched with deformable image registration (DIR) to a cone beam CT (CBCT) scan (b) which generates a deformation vector field (DVF) (c). The original contours of the pCT (d) are deformed according to the DVF (e) and copied to the CBCT (f).

1.6.4 Dose accumulation

DIR can facilitate dose reconstruction and subsequent accumulation of the reconstructed doses to assess the actually delivered dose to target volumes and OARs in the presence of anatomical changes [139]. One such dose reconstruction procedure matches the *source* image (pCT) to the *target* image (CBCT)

using DIR which results in a deformed pCT. This match is described by a '*forward*' DVF. The plan (beam arrangement, monitor units, fluence maps) is then recalculated on the deformed pCT. As deformed pCTs still encompass the correct Hounsfield unit (HU) calibration, they can be used to reconstruct the delivered dose. The resulting dose distribution is mapped back to the pCT space using the inverse or '*backward*' DVF. Reconstructed doses may then be accumulated in the pCT space. This method of dose accumulation is named all-to-one, where the 'one' implies a common reference frame. Figure 1.13 illustrates the all-to-one dose accumulation procedure for a pCT reference frame. Alternatively, a chronological accumulation of dose is possible where the reconstructed dose of a given fraction is mapped to the subsequent fraction's anatomy and summed with the current fraction's reconstructed dose. The resultant dose is then mapped to the following fraction's anatomy and the process is repeated. In the context of HNRT, a number of studies have utilised DIR-facilitated dose accumulation to evaluate the advantages of re-planning during the course of treatment [140,141], or to assess the impact of anatomical changes on delivered dose as a function of PTV margins [110,142]. However, comprehensive uncertainty analyses regarding accumulated dose are frequently lacking and when present, often do not sufficiently validate the accumulated dose [143]. Some studies have suggested direct calculation of the daily delivered dose in CBCTs (instead of dCTs) [144] by manipulating the CBCT HUs. However, intrinsic uncertainties in these adjustment techniques may confound the quantification.

1.6.5 Inverse consistency and symmetry

As introduced in *Section* 1.6, DIR is used to model the direction and magnitude of non-rigid tissue progression from different time-points during treatment. Accordingly, physically plausible DVFs that realistically represent the anatomical changes are desirable. To achieve these physically plausible DVFs, the DIR algorithm should be both symmetric and inverse consistent [139,145,146]. A symmetric algorithm that deforms the *source* image to the *target* image will produce a DVF that has the same transformation parameters but in opposite directions to the DVF produced when the *source* and *target* images are swapped. An inverse consistent algorithm derives a *source* to *target* DVF that is equal to

the mathematical inverse of the *target* to *source* DVF. Intuitively, the difference between symmetry and inverse consistency can be difficult to recognise as the two concepts are rarely separated in radiotherapy DIR literature. Furthermore, many contemporary DIR algorithms are bidirectional that guarantee inverse consistency by simultaneously optimising the '*forward*' and '*backward*' registrations [147–149]. Consequently, these contemporary DIR algorithms are also symmetric. The difference between symmetry and inverse consistency is shown in Figure 1.14 and becomes clearer when considering unidirectional DIR algorithms that only optimise the '*forward*' registration. Figure 1.14a illustrates the concept of symmetry. The '*forward*' *source* image to *target* image registration shown by the red solid arrow (Figure 1.14 left) is identical to the '*forward*' *target* image to *source* image registration shown by the blue solid arrow (Figure 1.14a right) **except** acting on opposite directions. Figure 1.14b illustrates inverse consistency. The *target* image to *source* image registration shown by the red dashed arrow (Figure 1.14b right) is the mathematical inverse of the *source* image to *target* image registration shown by the red solid arrow (Figure 1.14b left). Figure 1.14a shows symmetry and **not** inverse consistency. Figure 1.14b shows inverse consistency and **not** symmetry.

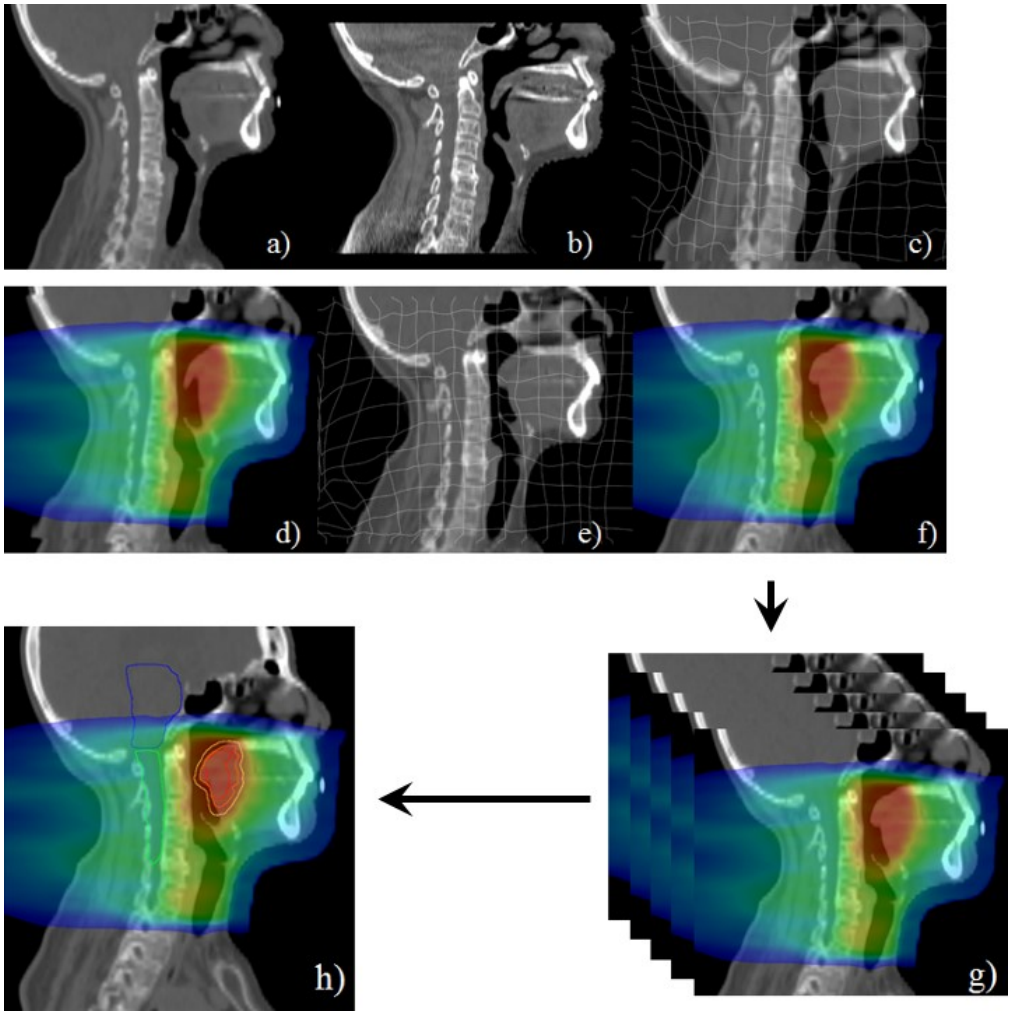


Figure 1.13: The all-to-one dose accumulation procedure. A planning computed tomography (pCT) scan (a) is matched to a cone beam CT (CBCT) scan (b) to generate a deformation vector field (DVF) and a deformed pCT (dCT) (c). The plan (beam arrangement, monitor units, fluence maps) is then recalculated (d) on the dCT. The inverse DVF (e) is applied to the recalculated dose and the resultant dose distribution is reconstructed in the pCT space (f). Reconstructed doses of successive fractions can be accumulated (g) in the pCT space (h).

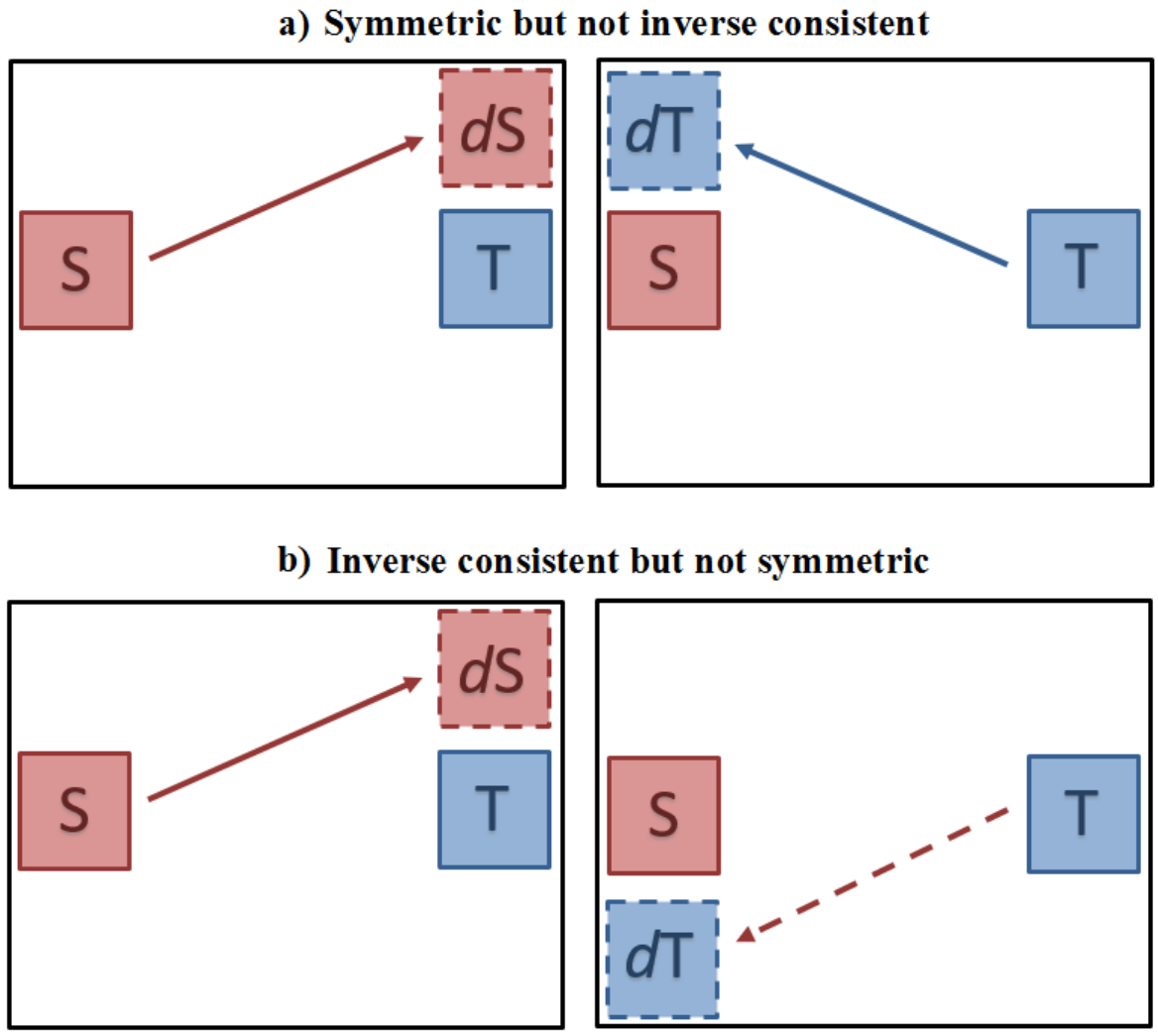


Figure 1.14: Two-dimensional schematic of the difference between symmetry and inverse consistency in deformable image registration (DIR) algorithms. (a, left) The source image (S) to target image (T) deformation vector field (DVF) is simplified and represented by the solid red arrow and results in a deformed source image (dS). (a, right) The T to S DVF is simplified and represented by the solid blue arrow and results in a deformed T (dT). (b, right) The T to S DVF is represented by the dashed red line and is the mathematical inverse of the solid red line. (a) symmetric but not inverse consistent considering that the solid lines are equal, opposite and not inverse consistent. (b) inverse consistent but not symmetric considering that the solid lines are equal, not opposite and inverse consistent.

1.6.6 DIR validation

It is imperative to validate any DIR software that is used in a clinical decision making process. However it must be stated that validation of DIR software is challenging because a comprehensive ground truth when dealing with patient data does not yet exist [130]. This is particularly problematic in regions without distinct tissue boundaries. Nevertheless, numerous studies have evaluated the accuracy and precision of DIR-facilitated contour propagation with the use of physician delineated landmarks and

contours [105,136–138,150–154]. Here, landmarks and contours delineated by the physician are typically compared to those deformed according to the DVF with a distance to agreement and/or overlap metric (or similar geometric analysis) [130]. Two commonly used approaches are the mean distance to agreement [155] and Dice similarity coefficient [156]. However considering the well documented intra- and inter-observer contour delineation variation [157,158] which confounds the physician established ground truth, many studies have instead opted to focus on the physical plausibility of the DVFs [131,159] by utilising metrics such as the Jacobian determinant of the DVF [130]. A number of studies have generated a ground truth deformation through use of a mechanical [160] or *in silico* phantoms [134,161] to validate their DIR software.

The methods described here so far (*Sub-section* 1.6.6) to validate DIR focus primarily on the *forward* DVF and associated applications, such as contour propagation. These validation methods are in agreement with guidelines published in the 2017 '*Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the American Association of Physicists in Medicine (AAPM) Radiation Therapy Committee Task Group No. 132*' report by Brock et al. [130].

The validation of DIR for dose accumulation was however outside the scope of the Brock et al. report [130] and in general, literature tackling this area is scarce. Compared to DIR-facilitated contour propagation where the accuracy is most important at the tissue boundaries, every voxel receiving clinically relevant dose is required to be accurately aligned for dose accumulation. Furthermore, the dose accumulation procedure described in *Sub-section* 1.6.4 is predicated on inverse consistent and symmetric '*forward*' and '*backward*' DVFs, as opposed to only an accurate '*forward*' DVF required in contour propagation. The potential for a displacement field error (DFE) in the '*forward*' registration caused by intrinsic limitations of DIR algorithms [134,153] as well as possible inverse consistency errors (ICEs) of the '*backward*' transformation [159] results in uncertainties in the accumulated dose [145,162–165]. A perfect '*forward*' and '*backward*' registration is shown for a simplified 2-dimensional case for a single voxel in Figure 1.15a. An incorrect '*forward*' registration caused by a DFE is illustrated in Figure 1.15b and an incorrect '*backward*' registration caused by an ICE is

illustrated in Figure 1.15c. As a consequence of DFEs and ICEs (which are themselves caused by a multitude of factors which are discussed in *Section* 3.4), the suitability of DIR for accurate and precise dose accumulation is debated in the literature [143,166–169]. To the best of the author's knowledge, only two studies have evaluated the uncertainty in dose accumulation for HNRT [145,162]. However, a comprehensive validation in these studies is hampered by not investigating the dose reconstruction uncertainties of structure sub-volumes or by not considering registration errors in the pseudo-ground truth registration.

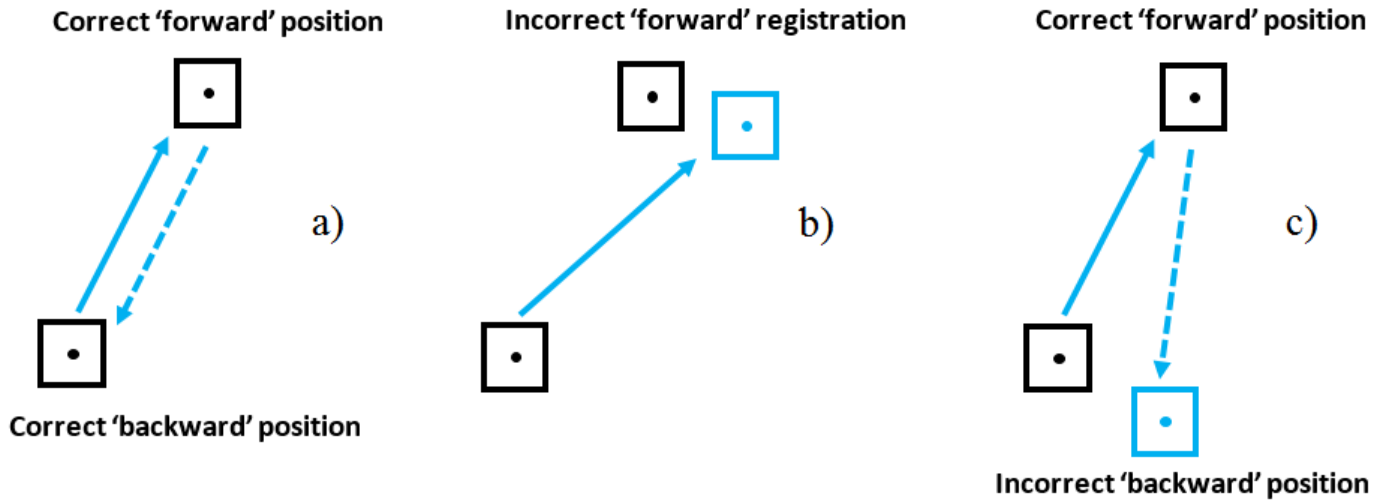


Figure 1.15: A simplified 2-dimensional case of a single voxel's 'forward' and 'backward' registration. (a) A perfect 'forward' and 'backward' registration shown by the voxel returning to its start position. (b) A displacement field error (DFE) is exemplified by the voxel's incorrect position after the 'forward' registration. An inverse consistency error (ICE) is exemplified by the voxel's incorrect position after the 'backward' registration.

1.7 The PTV margin in HNRT

As prefaced in *Section* 1.3, the PTV margin ensures adequate CTV coverage and accounts for various uncertainties including patient positioning, geometrical accuracy of the treatment machine and geometrical uncertainties of target volume definition. PTV margins and IGRT do not standardly account for the non-rigid deformations which are commonly observed in HNRT (*Sub-section* 1.4.2). Classical PTV margin recipe calculations [35,36] yield CTV to PTV margin expansions of 6-9 mm, even when non-rigid deformations are ignored and low estimates for the uncertainty in target delineation are used. Despite this, PTV margins of 5 mm are commonly applied to establish target coverage [170,171] and

high local tumour control rates up to 78-88% are commonly reported [1,2,53,62]. Most treatment failures occur well within the high-dose region, while recurrences within or near the PTV-margin boundary are exceptional [172–174].

1.7.1 PTV margin reduction

Three groups have reported on the clinical outcomes for patients treated with either 3 or 5 mm PTV margin plans [1–4]. Chen et al. [1,2] reported a 78% two-year LR control rate for both 3 and 5 mm PTV margin expansion patient groups (p = 0.96). Of the 46 (225 total) patients that had a loco-regional recurrence (LRR) in that study [2], eight were found to be marginal. A marginal LRR was observed in the dermal/subcutaneous surface, contralateral neck in the vicinity of spared PG and retropharyngeal area in three, three and two instances, respectively. Five and three of the eight marginal LRRs arose for patients treated with 5 and 3 mm PTV margin plans, respectively. It should be noted that these results may have been confounded by the application of adaptive radiotherapy (ART) for selected patients [119]. At three-year follow up [1], a significant reduction in oesophageal stricture (p = 0.01) and gastrostomy-tube dependence at one year ($p < 0.01$) was associated with a 3 mm PTV margin expansion. No significant differences in xerostomia or dysphagia were found for the different margins expansions.

Similarly, Navran et al. [3] did not observe a significant difference in the two-year LR control rates (p = 1.0) which were 79.9% and 79.2% for patients treated with 3 and 5 mm PTV margins, respectively. The geography of LRRs was not reported in that study. Navran et al. [3] applied an IGRT protocol that accounted for non-rigid patient positioning deviations as well as ART for selected patients. The median follow up time was 17.2 and 34.9 months ($p < 0.01$) for the 3 and 5 mm PTV margin groups, respectively. A significant reduction in acute grade three toxicity ($p = 0.03$), acute grade three mucositis ($p < 0.01$) and acute grade three dysphagia (feeding tube-dependence) (p = 0.03) was associated with the 3 mm PTV margin expansion patient group [3]. A significant reduction in the incidence of ongoing feeding tube-dependence after three months of radiotherapy was also observed for the 3 mm PTV margin patient group [3]. No significant differences in late grade ≥ two xerostomia were found.

In contrast to the previous two groups, Franzese et al. [4] found a significant difference ($p = 0.045$) in the two-year LR control rates which were 87.8% and 72.6% for patients treated with 5 and 3 mm PTV margins, respectively. Again, the location of LRRs was not reported. Daily IGRT was utilized for both margin expansion groups however it is unclear if an ART procedure was employed [4]. The median follow up time was 26 months. A significant reduction was not observed for any of the acute or late toxicity end points except for grade three acute dysphagia ($p < 0.01$) which, surprisingly, was lower for the 5 mm PTV margin patient group [4].

The difference between 3 and 5 mm PTV margins in visualised in Figure 1.16. It is important to optimise the PTV margin so appropriate target coverage is achieved while simultaneously ensuring the margins are not overly generous resulting in unnecessary toxicity for patients.

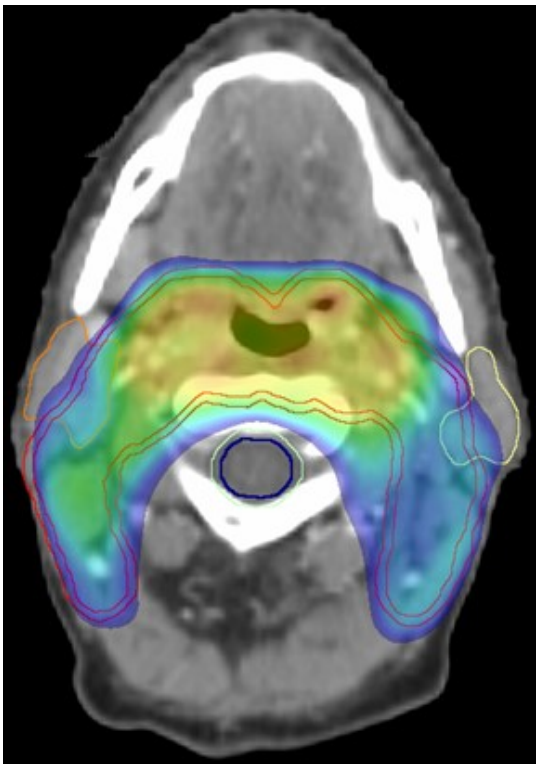


Figure 1.16: Visualisation of the difference between 3 (inner red contour) and 5 mm (outer red contour) planning target volume (PTV) margins. The colourwash represents the low-dose 95% isodose line. The parotid glands (orange and yellow contours) are located laterally. The spinal cord 3 mm planning risk volume (PRV) and 5 mm PRV are shown by the dark blue and light green contours, respectively.

1.8 Wellington Blood & Cancer Centre

Section 1.8 describes and compares WBCC practices with external departments in areas of HNRT that are pertinent to this dissertation's goal of developing a quality management "safety net" tool for patients who may be at risk of unacceptable loss in target coverage and/or overdosage to normal tissues with reduced PTV margins.

1.8.1 Contouring

Target delineation at the WBCC is carried out using pre-treatment CT imaging but is also based on clinical information and physical examination. Although magnetic resonance imaging (MRI) information is known to reduce inter/intra-observer variation [170,175], it is not always available for contouring at the WBCC. Furthermore, the potential added value of MRI imaging has not been exploited considering that patient positioning on MRI and CT is markedly different and DIR has not been implemented for MRI-pCT matching.

A reasonably common approach exists for target delineation although contouring is ultimately done at the discretion of the treating radiation oncologist and a strict contouring protocol does not yet exist for HNRT. A procedure for peer review of contours is currently being developed. Details of the WBCC target delineation approach for HNRT are:

- High-dose CTV (i.e., CTV_{66Gy}) = GTV + 5 mm margin + structures at risk. Bony anatomy is considered to be an anatomical barrier and is removed from the CTV (unless there is invasion). Air cavities are not consistently removed from the CTV
- Intermediate-dose CTV (i.e., CTV_{60Gy}) = intermediate risk structures; these are anatomical structures with a high probability of infiltration based on positron-emission tomography (PET) imaging and experience of the radiation oncologist
- Low-dose CTV (i.e., CTV_{54Gy}) = low risk structures as per The Danish Head and Neck Cancer Group (DAHANCA) consensus guidelines [176]

- High-, intermediate- and low-dose PTVs = high-, intermediate- and low-dose CTVs + 5 mm margin

The precision and reproducibility of target delineation at the WBCC has not been formally assessed. However, in a study by Ramadaan et al. [150] investigating the suitability of DIR for efficient contouring of structures upon treatment adaptation, it was found that DIR did offer a time saving in re-contouring [150]. The study also highlighted differences between independently redrawn contours and contours adapted after DIR, which was partially caused by intra-observer variability. These finding were consistent with those reported by other groups [152,177–179].

1.8.2 Accuracy of patient positioning

The introduction of CBCT for patient position verification at the WBCC in 2011 highlighted that the reproducibility of patient positioning, in particular the variation in the posture of the patient over the treatment course was larger than anticipated. This required a highly individualised patient positioning approach and posed a high workload for treatment staff to maintain adequate dose coverage, in particular when anatomical changes occurred during the treatment course. For these reasons, a multi-disciplinary group focussing on patient positioning was instigated in 2012. This group implemented various process changes based on the findings of weekly image review meetings. The weekly image review meetings were also started in 2012. A review of the patient positioning results using statistical process control (SPC) showed that the reproducibility of patient positioning significantly improved at the end of 2013 [85]. Furthermore, it was shown that the variation in positioning results from that point in time onwards could almost completely be attributed to sub-optimal positioning reproducibility for a limited number of cases (7.5%). CBCT guidance was initially carried out weekly in addition to daily IGRT using planar kilovolt (kV) imaging for HNRT. Daily CBCT guidance has been implemented since November 2015. At the WBCC, treatment couch corrections with three degrees of freedom (lateral, vertical, longitudinal shifts) are carried out before each treatment fraction, based on registration of the in-room CBCT imaging to the pCT.

It is difficult to directly compare the accuracy of patient positioning at the WBCC with other studies due to differences in measurement and analysis methods. However, the magnitude of patient deformation at WBCC [85] is considerably smaller than reported by Graff et al. [81], Polat et al. [82], van Kranen et al. [83] and Zhang et al. [84] (*Sub-section* 1.4.2). In a later study, van Kranen et al. [180] reported on strategies for patient position correction taking the difference in deviation of multiple regions of interest (ROIs) into account [180]. With these strategies, they succeeded in reducing the average residual deviation for the different ROIs and the overall impact of patient positioning errors. The residual deviations [180] were similar to those observed for most ROIs at the WBCC during patient positioning, assuming all patient position corrections at the WBCC are carried out based on the location of the reference ROI C1-C3. For the tilt of the patient's head, the results of van Kranen et al. [180] were significantly worse compared to the results of the WBCC. It is preferable to physically reduce a patient's non-rigid setup errors at pre-treatment setup, considering that treatment couch corrections only cannot reduce non-rigid patient positioning deformations for all ROIs to the same degree.

For the groups reporting on clinical outcomes after reducing PTV margins from 5 to 3 mm, the patient positioning accuracy achieved by the study of Navran et al [3] is that reported by the previously mentioned van Kranen et al papers [83,180]. The accuracy of patient positioning achieved by the group of Purdy et al. [1,2,52,119,181,182] has to the best of the author's knowledge not been documented in literature. In the previously described papers by the group of Purdy discussing the results of PTV margin reduction [1,2], it was stated that daily IGRT was applied including CBCT and image registration focussing on the high-dose PTV. This is similar to the current practice at the WBCC. However, this series of papers [1,2,52,119,181,182] also stated that standard head rests (TIMO, Med-Tec) for patient immobilisation were used in contrast to the individual head rests used at the WBCC. The individual head rests were fundamental for achieving a high accuracy in patient positioning. It is therefore likely that the group of Purdy et al. [1,2,52,119,181,182] did not achieve the same level of patient positioning accuracy as the WBCC while treating patients using reduced PTV margins. The accuracy of patient positioning achieved by the study of Franzese et al. [4,183] has also to the best of the author's knowledge not been documented in literature. In the work of Franzese et al. [4] it was stated that daily

CBCT-guided couch corrections were applied to correct patient positioning setup errors for patients treated with both 5 and 3 mm PTV margins.

1.8.3 Review of clinical results

A recent review of the WBCC clinical results of oropharyngeal cancer treatments found a non-significant difference between the three-year actuarial rates for LR control before and after the introduction of VMAT which were 87% and 89%, respectively [184]. These control rates are well in line with international literature reporting on LR control rates before and after the introduction of IMRT [53,62]. At follow-up, a significantly lower proportion of VMAT patients experienced grade two or higher acute dermatitis: 89% 3DCRT versus 55% VMAT ($p < 0.01$). At most time points, the proportion of VMAT patients experiencing grade 2 or higher late xerostomia was lower for VMAT than for 3D-CRT patients, although this difference was only statistically significant at 24 months after treatment ($p = 0.04$). The favourable toxicity profile could be related to a significantly lower dose to skin and parotid glands in VMAT patients. A PG D_{mean} less than 26 Gy is commonly used as the main objective to aim for to reduce xerostomia [25,59]. The study reviewing the WBCC clinical results [184] indicated that an additional future reduction of the PG D_{mean} by 5 to 10 Gy would result in a PG D_{mean} of less than 26 Gy for a considerably larger proportion of patients. Specifically, a D_{mean} less than 26 Gy for at least one of the PGs is estimated to be obtained for 47%, 63%, or 79% of the patients if either the current approach, a PG D_{mean} reduction of 5 Gy, or 10 Gy is pursued, respectively. These levels of PG D_{mean} reduction may be achieved by applying a PTV margin reduction from 5 to 3 mm.

1.8.4 Adaptive radiotherapy

Patients' non-rigid deformations are reviewed on an individual patient basis during weekly multidisciplinary team (MDT) image review sessions at the WBCC or by dedicated radiation therapists who review all patient treatment images. When these non-rigid changes are small and deemed to be clinically irrelevant, they are generally ignored. If these non-rigid changes are deemed to be of potential clinical relevance they are assessed by a radiation oncologist and in cases of contour and/or air cavity

changes, the physics department simulates the actually delivered dose using an HU override of the pCT. The non-rigid changes are then further discussed in an MDT setting and the necessity of a re-plan is debated considering multiple patient specific factors including whether a new CT would be representative for the remainder of treatment. If unaccounted for, these changes could cause geometrical miss and potentially result in under-dosage of tumour targets or over-dosage to OARs [103]. Changes that occur at the end of the treatment are often difficult to counteract, and re-planning is usually not possible due to time constraints. In these cases, avoiding an overdose to critical normal tissues is standardly prioritised over target coverage.

1.8.5 DIR

The clinical implementation of DIR at the WBCC is SmartAdapt (SA) [135] which is based on the demons algorithm [133,134] and is part of the commercial ARIA-radiotherapy management package (v13 Varian Medical Systems, Palo Alto, CA, USA). The following information is made available in Varian's SA user manual [135]:

The DIR algorithm used in SA is based on the original optical flow algorithm [133] with two primary modifications [134]. First, SA incorporated a coarse-to-fine multi-resolution approach to allow for relatively large deformation recovery without large errors. The original demons implementation assumed small deformations between image datasets and could not accurately recover large deformations. The coarse-to-fine multi-resolution approach begins by utilising low resolution images derived from the original datasets and as such, large deformations become relatively small which promotes convergence when solving the demons equation (*Sub-section* 1.6.2). After convergence, the 3D vectors of the displaced voxels (*Section* 1.6) are passed to the next higher resolution. The multi-resolution approach improves SA's efficiency relative to the original demons implementation [135]. Second, a symmetric force is built into SA. The symmetric force produces a transformation to map the source image to the target image consistent with the inverse transformation that maps the target image to the source image (Figure 1.14). Realistic patient non-rigid deformations are considered inverse

consistent and symmetric [130,131,146]. SA is limited in cases where multiple potential solutions exist. For example, a point inside an area of homogeneous image contrast on the *source* image could be linked to any other point in the homogeneous image contrast region on the *target* image. As such, sufficient information to describe the deformation may not be available in the image datasets. Another limitation of SA is if a voxel exists in the *source* image but not the *target* image. For example, a voxel of normal tissue may be present in the pCT but due to patient weight loss 'dissolves' from the CBCT. The SA DIR will not be able to find the one-to-one correspondence for this voxel. Therefore, the registration may not be accurate [135].

A small number of studies have evaluated the accuracy of SA propagated structures in the head-and-neck region utilising CT scans [104,138,150,185,186]. Investigations regarding the accuracy of structures deformed to CBCTs are even more limited [104,185,186].

1.9 Dissertation structure

This *Introduction* chapter presented a brief introduction to HNRT, DIR and current practices at the WBCC to provide context to this dissertation. As prefaced in *Section* 1.1 and detailed in *Sub-section* 1.7.1 , recent literature has reported reduced treatment-related toxicity after reducing PTV margins from 5 to 3 mm for HNRT [1–4]. However, equivalent LR control rates were not always maintained [4]. As it was considered whether it was possible to implement reduced PTV margins from 5 to 3 mm at the WBCC with the aim to improve patients' treatment-related toxicity, it was recognised that many department-specific aspects of HNRT including, but not limited to, plan design, constraint adoption, prioritisation of OARs and target coverage, technical aspects of treatment accuracy, robustness of the planning solution for anatomical changes and method of immobilisation can influence treatment outcomes [5] and should all be considered when the PTV margins are amended. Considering these sources of possible treatment variability, it is important to have adequate quality management tools in place particularly in cases of new treatment techniques that deliver much tighter dose distributions.

The overall goal of the research described in dissertation was to enable a safe PTV margin reduction in HNRT from 5 to 3 mm at the WBCC. One approach to safely implement a PTV margin reduction is to account for systematic changes in patient positioning and/or anatomy using timely and appropriate ART [115]. However, this approach requires quantification and separation of the different modes of non-rigid deformation as they can be corrected for with different adaptive methods. Furthermore, it requires investigating the dosimetric impact of non-rigid deformations. *Chapters Two, Three* and *Four* detail the development of the tools to (*a*) quantify and separate the different modes of non-rigid deformation; and (*b*) investigate the dosimetric impact of the non-rigid deformations. These tools aimed to form the quality management "safety net" to detect those patients who may be at risk of unacceptable loss in target coverage and/or overdosage to OARs with reduced PTV margins.

Chapter Two introduces Exponentially Weighted Moving Average (EWMA) Statistical Process Control (SPC) charts [187] in combination with DIR to monitor and quantify individual patient non-rigid deformations during treatment. An individual patient's positioning reproducibility and non-rigid deformations must be well controlled and monitored to enable safe PTV margin reduction. SPC is widely used in manufacturing as a quality management tool and recently, its use in radiotherapy for quality assurance has been discussed [188]. EWMA charts, also known as process-behaviour charts, are frequently used as an SPC tool to examine and monitor process time trends [189]. The unambiguous instructions surrounding control chart initialisation and interpretation allow time trends to be reviewed in a standardised way. To the best of the author's knowledge, no previous study has reported on trend analysis that quantifies and separates different modes of patient-specific anatomical change using SPC and DIR.

Chapters Three and *Four* are closely related to each other considering their predication on dose reconstruction and accumulation. Dose reconstruction and accumulation (as introduced in *Sub-section* 1.6.4) in the context of the research described in this dissertation refers to the utilisation of DIR-facilitated reconstruction and accumulation of the delivered dose between different treatment time points. Consequently, the actually delivered dose for a HNRT treatment in the presence of anatomical

changes can be quantified. The dose accumulation procedure may be particularly applicable to the objective assessment of the adequacy of target coverage during HNRT with reduced PTV margins. However, registration errors (*Sub-section* 1.6.6) can introduce an uncertainty in the dose reconstruction and accumulation [145,162–164]. *Chapter Three* utilised an *in silico* ground truth based on clinically observed deformations to investigate the uncertainty of a DIR-facilitated dose accumulation workflow. The uncertainty in the accumulated dose was incorporated as a confidence interval in DVHs. Dose accumulation is currently not facilitated in Varian's SA module and requires external software. Considering that only limited literature is available on the accuracy of DIR-facilitated dose accumulation in general and that no previous study has reported on the accuracy of the intended clinical workflow, the retrospective study in *Chapter Three* was instigated.

Chapter Four investigated the impact of anatomical changes on the robustness of treatment plans constructed with 5 and 3 mm PTV and PRV margins. DIR-facilitated dose accumulation was used to investigate the robustness of treatment plans' target coverage, critical OAR avoidance and non-critical OAR sparing. The target coverage, critical OAR avoidance and non-critical OAR sparing was first quantified with DVH metrics. However, DVH metrics inherently lack spatial information regarding the dose distribution within a volume of interest [190] and for large target volumes, a clinically relevant loss in local sub-volume coverage may not be detected. Therefore, a more in-depth voxel-specific investigation was conducted. In the context of reducing PTV margins from 5 to 3 mm in HNRT, the retrospective study in *Chapter Four* contributed to the scarce body of literature regarding the robustness of treatment plans for anatomical changes and presented key recommendations.

Chapters Five and *Six* contain the *Discussion* and *Conclusion*, respectively. The *Discussion* chapter presents further considerations that are relevant to each study which are not provided in the corresponding discussion *Sections* (i.e., 2.4, 3.4 and 4.4) and links the individual studies together to present a collective summary showing how they constitute a coherent body of research. Conclusions to be taken from the research in this dissertation are presented in *Chapter Six*.

2 MONITORING ANATOMICAL CHANGES OF INDIVIDUAL PATIENTS USING STATISTICAL PROCESS CONTROL DURING HEAD-AND-NECK RADIOTHERAPY

Full citation,

- Lowther, N.J., Hamilton, D.A., Kim, H., Evans, J.M., Marsh, S.H. and Louwe, R.J., 2019. Monitoring anatomical changes of individual patients using statistical process control during head-and-neck radiotherapy. *Physics and Imaging in Radiation Oncology, 9*, pp.21-27

Author contributions,

I, along with co-author RL developed the strategy of monitoring individual patients' anatomical changes with SPC and DIR. I collected all patient data, performed all tasks within the treatment planning system and wrote all the MATLAB scripts for data analysis. I integrated standardised uncertainty equations with the mathematics of SPC and performed all the data and statistical analyses. Co-authors DH, HK and JE are consultant radiation oncologists at the Wellington Blood & Cancer Centre and contributed to this study by providing re-contouring of the PGs on CBCTs. I, along with co-authors SM and RL developed the conclusions. I wrote this paper with modifications to structure, flow and readability from SM and RL.

Monitoring anatomical changes of individual patients using statistical process control during head-and-neck radiotherapy

Nicholas J. Lowther[1,2], David A. Hamilton[1], Han Kim[1], Jamie M. Evans[1], Steven H. Marsh[2], Robert J. W. Louwe[1]

[1]*Wellington Blood and Cancer Centre, Department of Radiation Oncology, Wellington, New Zealand*

[2]*University of Canterbury, School of Physical and Chemical Sciences, Christchurch, New Zealand*

Keywords: anatomical changes, head-and-neck, radiotherapy, statistical process control, deformable image registration

Abstract

Background and purpose: Reduced toxicity while maintaining loco-regional control rates have been reported after reducing planning target volume (PTV) margins for head-and-neck radiotherapy (HNRT). In this context, quantifying anatomical changes to monitor patient treatment is preferred. This retrospective feasibility study investigated the application of deformable image registration (DIR) and Exponentially Weighted Moving Average (EWMA) Statistical Process Control (SPC) charts for this purpose.

Materials and methods: DIR between the computed tomography for treatment planning (pCT) images of twelve patients and their daily on-treatment cone beam computed tomography (CBCT) images quantified anatomical changes during treatment. EWMA charts investigated corresponding trends. Uncertainty analysis provided 90% confidence limits which were used to confirm whether a trend previously breached a threshold.

Results:

Trends in patient positioning reproducibility occurred before the end of treatment week four in 54% of cases. Using SPC process limits, only 24% of these were confirmed at a 90% confidence level before the end of treatment. Using an *a priori* clinical limit of 2 mm, absolute changes in patient pose were detected in 39% of cases, of which 82% were confirmed. Soft tissue trends outside SPC process limits occurring before the end of treatment week four were confirmed in 90% of cases.

Conclusion: Structure specific action thresholds enabled detection of systematic anatomical changes during the first four weeks of treatment. Investigation of the dosimetric impact of the observed deviations is needed to show the efficacy of SPC to timely indicate required treatment adaptation and provide a safety net for PTV margin reduction.

2.1 Introduction

The introduction of intensity modulated radiotherapy has enabled highly conformal dose deliveries which allows dose reduction to organs at risk (OAR) and reduced treatment toxicity [21,191,192]. These highly conformal treatments require image-guided radiotherapy (IGRT) to warrant accurate patient positioning and monitoring of changes in patient anatomy [115]. Planning target volume (PTV) margins of 5 mm are commonly applied to establish target coverage [170,171]. PTV margins and IGRT do not standardly account for non-rigid anatomy changes which are commonly observed in head-and-neck radiotherapy (HNRT) [78]. Nevertheless, two retrospective studies have suggested that 3 mm planning target volume (PTV) margins are sufficient if daily IGRT is used to correct for rigid patient position variations [193,194]. The study of Chen *et al.* has also clinically implemented reduced PTV margins from 5 mm to 3 mm and reported a reduction in late toxicity side-effects while maintaining equivalent two- and three-year loco-regional control rates [1,2,194]. However, these results may have been confounded by the application of adaptive radiotherapy (ART) for selected patients [119]. A recent study by Navran *et al.* also reported on favorable toxicity profiles for reduced 3 mm PTV margins while maintaining good tumor control rates [3]. The latter study applied an IGRT protocol that accounted for non-rigid patient positioning deviations as well as ART for selected patients.

As we considered whether it was possible to implement reduced 3 mm PTV margins after achieving a large improvement in patient positioning in our department [85], it was recognized that many other aspects of HNRT including accuracy of target delineation, robustness of the planning solution for anatomical changes influence treatment outcomes [5] and should all be considered when the PTV margins are reduced. In addition, objective and generally applicable guidelines to select patients who will benefit from treatment adaptation are lacking [79,115], and most studies on ART apply subjective criteria to select patients for treatment adaptation [93,107,119,121]. In an effort to create a safety net for patients who exhibit large non-rigid deformations and objectively select patients who would benefit from treatment adaptation in the context of PTV margin reduction, this feasibility study investigated the first step to build such a framework. Specifically, the suitability of deformable image registration (DIR) and exponentially weighted moving average (EWMA) statistical process control (SPC) charts to

objectively quantify and monitor individual patients' deformation, i.e., non-rigid changes in both pose and anatomy of the patient during treatment were investigated.

2.2 Materials and methods

2.2.1 Patient group

This retrospective study included twelve patients with cancers of the head-and-neck (HN) region that were treated radically. Comprehensive information regarding the patient cohort can be found in the Supplementary Data (**A.1:** Patient group). The patients included in this retrospective study provided written consent to use their data in clinical audits.

2.2.2 Planning CT contouring

Nine bony anatomy (BA) structures were delineated on each patient's computed tomography for treatment planning (pCT) scan: C1-C3, C3-C5, C5-C7, mandible, maxilla, base of skull, hyoid, occipital and larynx. A pre-defined neck volume V^{neck} was defined as all tissue within the volume bounded by axial planes at the base of C3 at the anterior cortical boundary and base of C4 at the anterior cortical boundary (Figure 2.1). The original parotid gland (PG) contours were reviewed by a radiation oncologist and corrected where necessary to minimize the uncertainty of the DIR results.

2.2.3 Deformable image registration

The pCT of each patient was deformed to match the anatomy of each daily cone beam computed tomography (CBCT) using SmartAdapt v.13 (Varian Medical Systems, Palo Alto CA, USA), resulting in 30 deformed CTs (dCT). The deformed structures of the dCT as well as the non-deformed structures of the pCT were propagated to the corresponding CBCTs. These data sets were exported in Digital Imaging and Communications in Medicine (DICOM) format for subsequent analysis with in-house developed software (MATLAB 2016b version 9.1, The MathWorks Inc.).

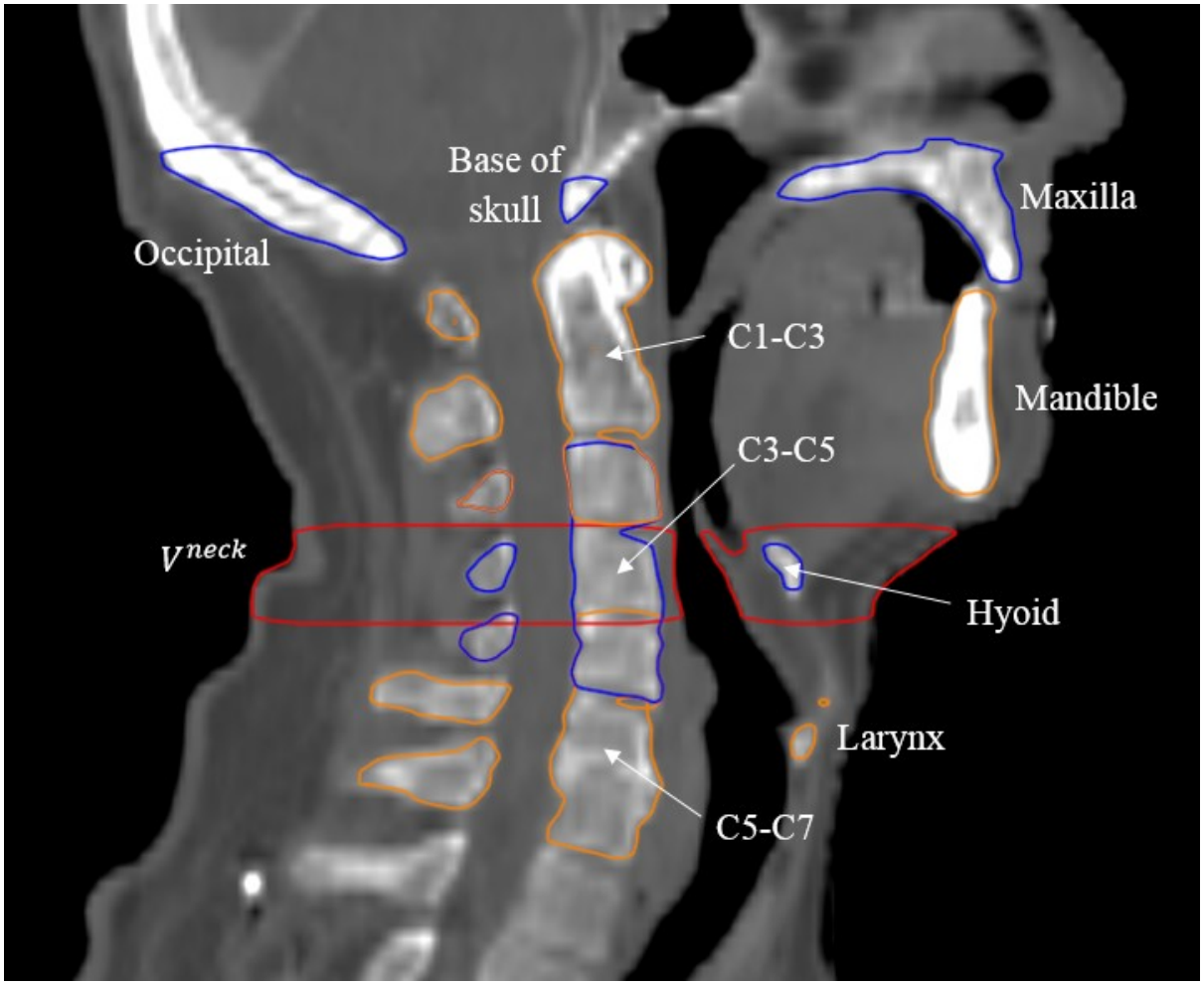


Figure 2.1: Definition of bony anatomy (BA) structures and neck volume V^{neck} in the central sagittal plane. The superior end of the occipital bone contour was defined as to just include the external occipital protuberance.

2.2.4 Quantifying non-rigid deformations

The exported contours were used to quantify the non-rigid deformations: BA positioning reproducibility, centroid shifts of the PGs, volume changes of the PGs (V^{PG}) and V^{neck} changes. The BA positioning reproducibility was quantified by first calculating the change T in centroid position C of all BA structures with respect to its pCT position for each direction k = x, y, z, and each fraction f:

$$T_{f,k}^{BA} = C_{f,k,dCT}^{BA} - C_{f,k,pCT}^{BA} \quad (1)$$

Subsequently, the 3D-deviations of all BA structures were calculated relative to the reference structure C1-C3, using:

$$\partial_f^{BA} = \sqrt{\left(T_{x,f}^{BA} - T_{x,f}^{C1-C3}\right)^2 + \left(T_{y,f}^{BA} - T_{y,f}^{C1-C3}\right)^2 + \left(T_{z,f}^{BA} - T_{z,f}^{C1-C3}\right)^2} \quad (2)$$

Similarly, the 3D-shifts of the PGs were calculated relative to C1-C3 according to Eq. 2.

To assist with readability, the 3D-deviation of BA structure X relative to C1-C3 will be referred to as "deviation of X" from here onwards.

2.2.5 Statistical process control

The trends of the metrics ∂_f^{BA}, ∂_f^{PG}, V_f^{neck}, V_f^{PG} over time were monitored using EWMA charts [195]. The values E_f of the EWMA statistic were calculated using:

$$E_f = \lambda X_f + (1 - \lambda) E_{f-1} \quad (3)$$

where λ is a constant defined as $0 < \lambda \leq 1$ that determines the depth of memory (smoothing) of the EWMA and X_f is the metric of interest for fraction *f*. As shown in Figure 2.2, E_0 is the average metric of interest during the EWMA reference period. The first week of treatment (i.e. fractions 1-5) is used as the EWMA reference period, assuming that prospective monitoring of EWMA statistics would be conducted from fraction six onwards in a clinical scenario, with EWMA statistics for fractions 1-5 analyzed retrospectively.

The lower and upper process limits LPL and UPL, which indicate a statistically significant difference from the reference period, were calculated using:

$$\text{LPL, UPL} = \mu_0 \pm L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left[1 - (1-\lambda)^{2f}\right]} \quad (4)$$

The initial deviation μ_0 and variation σ were calculated as the mean and SD for the reference period, and *L* is a factor determining the width of the process limits. Considering that Q-Q plots indicated the

metrics of interest were not always normally distributed, λ and L were set to $\lambda = 0.05$ and $L = 2.492$ to obtain similar type I and type II error probabilities as for normally distributed data [187].

PGs were grouped according to mean dose at planning: low- (< 26 Gy, n = 7) and high-dose (≥ 26 Gy, n = 17) consistent with the different impact of radiation dose predicted below or above this cut-off point by normal tissue complication probability (NTCP) models [25,59]. In cases where the superior or inferior extents of the CBCT did not contain a specific BA structure, that structure was excluded from the analysis for that patient.

The fraction numbers at which the EWMA statistics went outside the corresponding process limits were recorded. Where possible, EWMA trends that were obtained using DIR were compared with results obtained using an alternative manual method. ∂_f^{BA} results obtained using DIR were compared with results obtained using a manual method of registration (MMR) as described by various authors [83,85]. The MMR used rigid image registrations (translations only; no rotations) of the individual structures to assess the shift of the structure centroids relative to the pCT. DIR results for V_f^{neck} trends were compared with the change of a single axial slice volume at the base of C2 that was manually contoured for all CBCTs as well as changes in patient weight. Considering that a change in patient weight may occur at any location of the patient's body, V_f^{neck} trends were expected to describe volume change of the treatment region more appropriately than a simple weight metric.

2.2.6 Detection of trends

In this study two methods to detect trends were investigated:

2.2.6.1 SPC limits for trend detection

A detailed uncertainty analysis was carried out to assess whether DIR and EWMA charts can accurately monitor non-rigid deformation trends of individual patients, and detect changes in patient positioning or anatomy in a timely manner. In these analyses including both the DIR and manual results, the uncertainty was calculated as per the Guide to the Expression of Uncertainty in Measurement (GUM) [196] (see Supplementary Data **A.2:** Uncertainty analysis). The results of the uncertainty analyses were used to quantify the accuracy of trends as the average 90% level of confidence interval of the EWMA statistics for each metric over all treatment fractions. The sensitivity to detect a trend (i.e., the minimal detectable trend magnitude) with 90% confidence with respect to the reference period (i.e., E_0 set to zero) was defined as the difference between the 90% level of confidence of the upper EWMA process limit and the EWMA center line at the last treatment fraction. Differences between the sensitivity and accuracy of the two methods over all patients were assessed using paired *t*-tests at the 0.05 level of significance. The robustness of trend detection using SPC limits was quantified as the proportion of cases where the accuracy interval of the trend exceeded the 90% level of confidence of the process limits before the end of treatment relative to the number of cases where the EWMA statistic itself exceeded a process limit before the end of the fourth week of treatment.

2.2.6.2 Clinical limit for trend detection

Preliminary SPC trend results highlighted sub-mm process limits for specific structures which would likely not provide an efficient threshold for clinically relevant process changes. Also the exact moment that a BA deviation trend exceeded the control limits within a certain statistical confidence level was not easily defined, in particular for shallow trends in combination with narrow process limits of a structure. The EWMA trend analysis was therefore also carried out using an *a priori* 2 mm clinical limit as detection threshold for *absolute* BA deviations. The robustness of trend detection using a 2 mm clinical limit (including the observed E_0) was defined as the proportion of cases where the accuracy

interval of the EWMA statistic exceeded this limit relative to the number of cases where the EWMA statistic itself exceeded this clinical limit before the end of the fourth week of treatment.

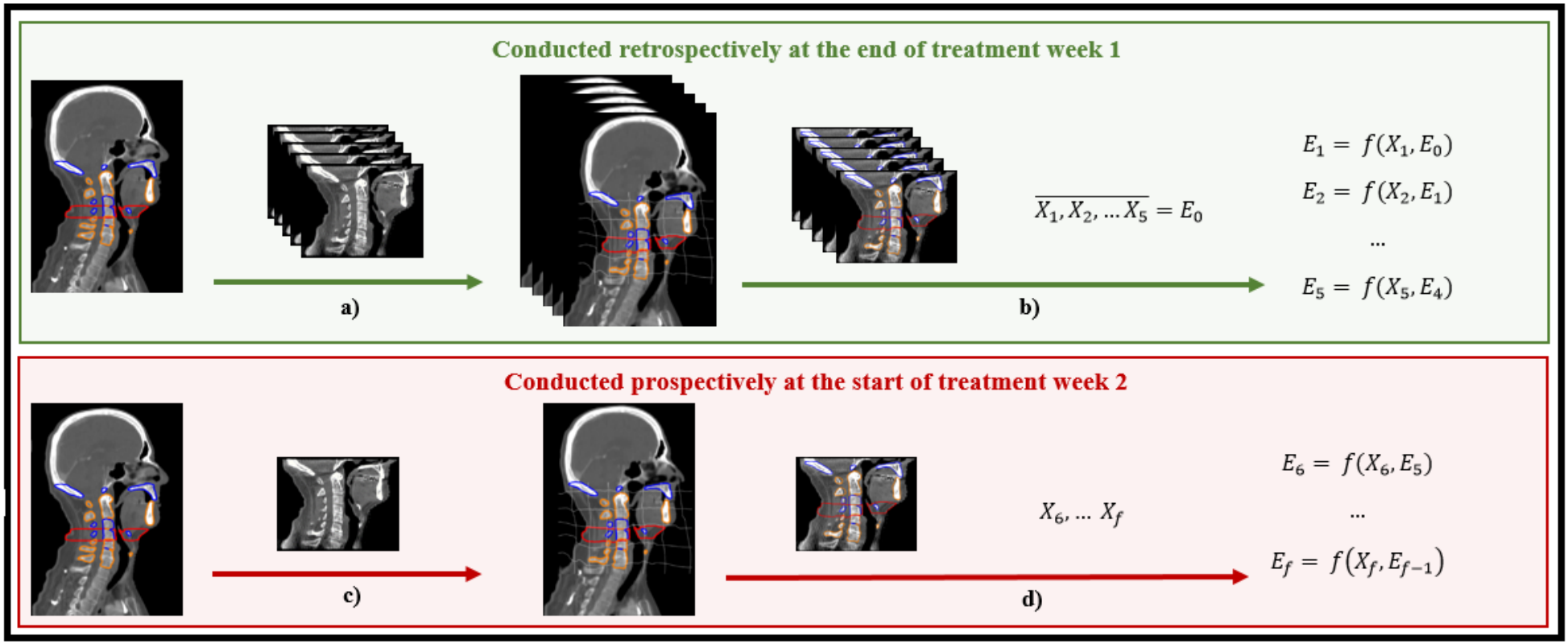


Figure 2.2: Workflow detailing the exponentially weighted moving average (EWMA) statistics monitoring procedure based on DIR. (a) The planning CT (pCT) is deformed to match the anatomical configurations of each cone beam CT (CBCT) of treatment week 1; (b) Deformed pCT contours propagated to the CBCTs are used to calculate metrics of interest X_f for fractions 1-5. The average metric of interest is defined as E_0 and used to calculate the EWMA statistics $E_1 \ldots E_5$ as per Eq. (3); (c-d) At the beginning of treatment week two, propagation of deformed contours and EWMA calculations are conducted on a fraction-by-fraction basis.

2.3 Results

2.3.1 SPC trends

Overall, 31% of the trends describing BA positioning reproducibility acquired with DIR remained within the EWMA process limits during treatment, while 54% and 15% of the trends exceeded a process limit before and after the end of the fourth week of treatment (i.e., fraction 20), respectively (Supplementary Data **A.3:** SPC results). Figure 2.3a shows the deviation of the mandible for patient three exceeding the upper DIR process limit at fraction 18.

2.3.1.1 SPC limits for trend detection

Both the average accuracy and sensitivity calculated over all patients were smaller than 1 mm for BA deviations. Table 2 summarizes the average accuracy and sensitivity recorded for the various structures and compares the results obtained for DIR and MMR. Except for the deviation of the occipital bone calculated with DIR, the average accuracies were smaller than the corresponding average sensitivities indicating that the minimal detectable trend with a 90% level of confidence was generally equal to the observed sensitivity. Statistically significant differences between DIR and MMR accuracies were observed for all BA structures except deviation of the maxilla. Conversely, only the sensitivity to detect a C5-C7 deviation was significantly different between the two methods.

The robustness to detect a statistically significant change relative to the reference period for each trend is summarized in Table 3 (individual patient results available in Supplementary Data **A.4:** Individual patient 90% level of confidence SPC results). For BA deviations, the overall robustness of trend detection was 24% (95% CI: 11-38%), with none of the potential trends for occipital, base of skull, maxilla and larynx deviations confirmed to be an actual trend at a 90% level of confidence. In contrast, the robustness of trend detection for soft tissue trends was 90% (95% CI: 80-99%), due to steeper gradients of these trends.

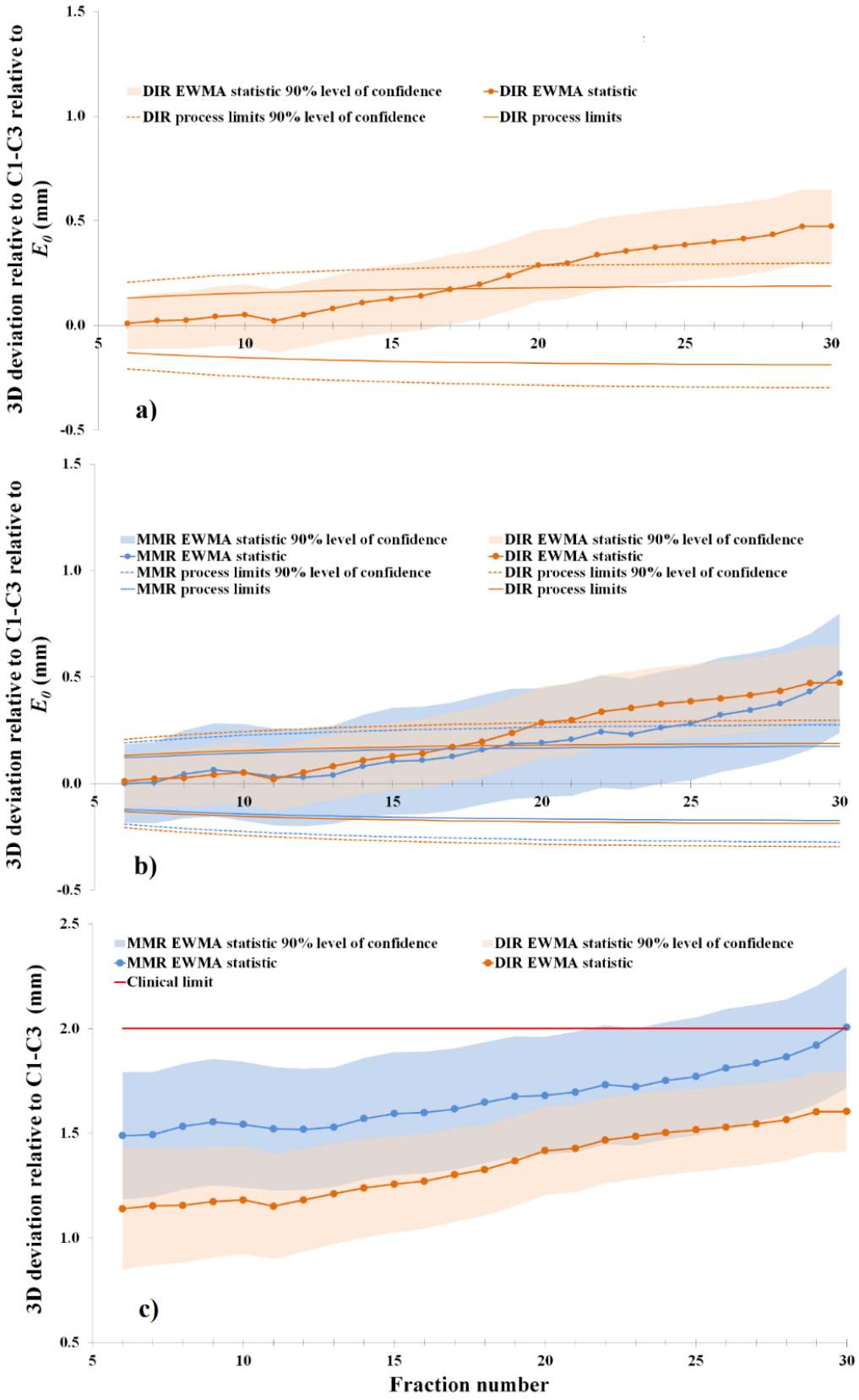


Figure 2.3: Example of the exponentially weighted moving average (EWMA) charts used in this study of the mandible deviation for patient three. EWMA trends derived from deformable image registration (DIR) or the manual method of registration (MMR), are shown in orange and blue, respectively. Trends are plotted either as deviations relative to the SPC reference period (panels a and b) or absolute deviations including E_0 (panel c). (a) DIR derived EWMA chart including uncertainty analysis and corrected for initial deformation E_0; (b) comparison of DIR and MMR derived EWMA charts; (c) DIR and MMR derived EWMA charts including observed E_0 for the purpose of comparing the trends with a 2 mm clinical limit.

2.3.1.2 *Clinical limit for trend detection*

The *absolute* BA deviations exceeded the 2 mm clinical limit in 33/84 (39%) of the available cases (Table 4) (individual patient results available in Supplementary Data **A.5**: Individual patient 90% level of confidence clinical limit results). Figure 2.3c provides an example of this analysis where the deviation of the mandible for patient 3 acquired using DIR did not exceed the 2 mm clinical limit during the treatment period. In the majority of cases (29/33) these deviations were already larger than 2 mm from the start of treatment in particular for the hyoid and larynx, indicative of the high mobility of these structures (see also Supplementary Data **A.6:** Differences between E0 values of SPC charts based on DIR or MMR**)**. The overall robustness of trend detection for BA deviations larger than 2 mm was 82% (95% CI: 67-96%).

2.3.1.3 *Trends returning to control*

EWMA trends for BA deviations did not move back within either the process limits or clinical limits in any of the cases after previously exceeding control at the 90% level of confidence. In 2 - 3% of cases a prior confirmation of a trend was negated as the lower boundary of the trend's accuracy interval moved back either within the upper process limit's 90% level of confidence or the clinical limit.

2.3.2 Soft tissue analysis

For eight patients (73%), the trend describing a volume change of V_f^{neck} exceeded a process limit before the end of the fourth week of treatment, and for one patient (9%), this occurred during the final two weeks of treatment. These trends usually represented a decrease of the neck volume, however for two of these nine patients, an increase in V_f^{neck} was observed.

A strong correlation ($r^2 = 0.87$; $p < 0.01$) was found between the changes in V_f^{neck} and patient weight. A slightly lower correlation ($r^2 = 0.76$; $p < 0.01$) was found between the changes in V_f^{neck} and those of the single axial slice volume at the base of C2. The 90% confidence intervals of the EWMA trends of the latter two metrics overlapped in 91% of cases.

Ninety-four percent of the 48 trends describing shift or shrinkage of a PG exceeded a process limit during HNRT treatment. EWMA trends describing shrinkage of high and low mean dose PG groups exceeded a process limit before the end of the fourth week of treatment in 94% and 86% of the cases, respectively. A decrease in volume was observed for all parotid glands. EWMA trends describing PG shift exceeded a process limit before the end of the fourth week of treatment in 76% and 71% of the cases, respectively. The Supplementary Data (**A.7:** Average parotid shift and volume EWMA) shows EWMA trends of average volume and shift for both PG groups.

The average volume change of the PGs from planning to end of treatment was not significantly different ($p=0.20$) for the high dose PGs (-5.9 cm^3; range -3.3 to -10.0 cm^3) and the low dose PGs (-4.8 cm^3; range: -3.5 to -9.4 cm^3). Similarly, the average shift during treatment was not significantly different ($p=0.38$) for the high dose PGs (2.1 mm; range 1.1 to 3.1 mm) and the low dose PGs (1.1 mm; range 1.1 to 3.3 mm).

*Table 2: Average sensitivity S and accuracy A of exponentially weighted moving average (EWMA) charts established from deformable image registration (DIR) and manually acquired raw data (MMR). Δ represents the difference between DIR and MMR results. Accuracy A= average 90% level of confidence interval over all treatment fractions. Sensitivity S= half the EWMA process limit 90% level of confidence at the last treatment fraction. *indicates a statistically significant difference.*

	DIR		**MMR**		**Δ**	
Structure	***A***	***S***	***A***	***S***	***A***	***S***
Occipital bone [mm]	0.3	0.3	0.2	0.3	0.1*	0.0
Base of skull [mm]	0.2	0.3	0.2	0.3	0.0*	0.0
Maxilla [mm]	0.2	0.3	0.2	0.4	0.0	-0.1
C3-C5 [mm]	0.1	0.3	0.2	0.3	-0.1*	0.0
C5-C7 [mm]	0.1	0.4	0.2	0.5	-0.1*	-0.1*
Hyoid [mm]	0.2	0.7	0.3	0.7	-0.1*	0.0
Larynx [mm]	0.2	0.8	0.3	0.8	-0.1*	0.0
Mandible [mm]	0.2	0.4	0.2	0.4	0.0*	0.0
High dose PG dev. [mm]	0.1	0.3	-	-	-	-
Low dose PG dev. [mm]	0.1	0.3	-	-	-	-
High dose PG vol. [cm^3]	0.3	0.5	-	-	-	-
Low dose PG vol. [cm^3]	0.3	0.5	-	-	-	-
V_f^{neck} [cm^3]	2.0	3.9	-	-	-	-

Table 3: Median and range of the fraction number where the exponentially weighted moving average (EWMA) statistic exceeded an SPC process limit before the end of treatment week 4, and the proportion of cases where the existence of a trend was confirmed by the accuracy interval of the trend exceeding the process limits 90% level of confidence before the end of treatment. This proportion is also expressed as percentage representing the robustness of the trend detection with the corresponding 95% confidence interval. EWMA parameters were established from deformable image registration (DIR) raw data. (PG = parotid gland; BOS = base of skull; Occ = occipital; Max = maxilla; Mand = mandible).

	∂_f^{Occ}	∂_f^{BOS}	∂_f^{Max}	∂_f^{C3-C5}	∂_f^{C5-C7}	∂_f^{Hyoid}	∂_f^{Larynx}	∂_f^{Mand}	$\partial_f^{PG\ high}$	$\partial_f^{PG\ low}$	$V_f^{PG\ high}$	$V_f^{PG\ low}$	V_f^{neck}
Median	13	15	11	14	14	14	19	15	14	9	10	12	13
Range	12-19	9-20	8-15	10 -18	12-19	11-17	14-20	8-18	9-20	7-18	8-20	9-20	9-18
Proportion confirmed	0/4	0/6	0/5	3/7	3/5	2/7	0/4	3/7	12/13	4/5	14/16	6/6	7/8
Robustness	0%	0%	0%	43%	60%	29%	0%	43%	92%	80%	88%	100%	88%
95% CI	0-13%	0-8%	0-10%	0-87%	7-100%	0 - 69%	0 -13%	0% -87%	74-100%	35-100%	68-100%	92-100%	58-100%

Table 4: Median and range of the fraction number where the exponentially weighted moving average (EWMA) statistic exceeded a 2 mm clinical limit before the end of treatment week 4, and the proportion of cases where the existence of a trend was confirmed by the accuracy interval of the trend exceeding the clinical limit before the end of treatment. This proportion is also expressed as percentage representing the robustness of the trend detection with the corresponding 95% confidence interval. EWMA parameters were established from deformable image registration (DIR) raw data. (BOS = base of skull; Occ = occipital; Max = maxilla; Mand = mandible).

	∂_f^{Occ}	∂_f^{BOS}	∂_f^{Max}	∂_f^{C3-C5}	∂_f^{C5-C7}	∂_f^{Hyoid}	∂_f^{Larynx}	∂_f^{Mand}
Median	1	1	6	NA	1	1	1	1
Range	1-1	1	1-11	NA	1-19	1	1	1-10
Proportion confirmed	3/4	0/1	1/2	0/0	3/3	9/9	8/9	3/5
Robustness	75%	0%	50%	NA	100%	100%	89%	60%
95% CI	20 -100%	0 -50%	0 -100%	NA	83 -100%	94 -100%	63 -100%	7 -100%

2.4 Discussion

This study investigated the suitability of DIR and EWMA SPC to quantify and monitor individual patients' changes in pose and anatomy during HNRT. This method facilitates a standardized approach to quantify patient deformations with an estimated robustness derived from comprehensive uncertainty analyses. It is a first step in developing a safety net for PTV margin reduction in HNRT as well as objective guidelines to select patient for treatment adaptation. Considering that loss in target coverage during treatment might occur more often and might become more relevant with reduced mm PTV margins, EWMA-facilitated detection of non-rigid changes in pose and anatomy during treatment of an individual patient could allow for early treatment adaption.

Previous studies investigated the application of SPC charts to monitor the reproducibility of patient positioning based on MMR [85]. However, MMR is very labour and time intensive; hampering clinical implementation. The application of DIR and SPC to quantify and monitor individual patient deformations during treatment can largely be automated and is therefore attractive from an economic and efficiency perspective. Considering that a change in treatment is usually not feasible in the final weeks of treatment due to the time required for re-scanning the patient, contouring and re-planning, this study assumed that clinically relevant changes in pose and anatomy should be detected before the end of treatment week four.

Over all patients, DIR EWMA trends exceeded a process limit before the end of treatment week 4 in 54%, 73% and 83% of cases for BA deviations, volume changes of V_f^{neck} and PG changes, respectively (Supplementary Data **A.3:** SPC results). However, these process limits represent the boundaries of the variation expected for a 'process' based on the observed variation during the reference period. The SPC limits therefore ignore the offset E_0 during the reference period relative to the desired patient position as defined during CT-planning which can be several mm's in case of highly mobile structures. Therefore, deformation trends exceeding an SPC process limit itself is not indicative for the clinical relevance of an observed BA deviation and would require additional interpretation rules. In addition, potential trends occurring during the reference period add to the uncertainty of the SPC limits. Furthermore, the exact moment that the trend of a BA structure deviation relative to the initial value during the reference period exceeded an SPC process limit was not easily defined due to the shallow

trends that were observed, in spite of the sub-mm accuracies of the EWMA statistic and process limits. The latter problem is exemplified in Figure 2.3a where the EWMA statistic exceeds the upper process limit at fraction 18, but the 90% level of confidence interval of the EWMA statistic does not cross the 90% level of confidence of the upper process limit during the treatment course. Similar behavior was observed for MMR derived EWMA charts (Figure 2.3b). At the 90% level of confidence, BA deviation trends where the EWMA statistic exceeded the process limits during the first four weeks of treatment were only confirmed in 24% of the cases. For these reasons, it is more efficient to apply a clinical limit to detect BA deviations that could potentially have a clinical impact during treatment. This study showed that EWMA statistics representing BA deviation trends exceeding a generic *a priori* 2 mm clinical limit before the end of treatment week four were detected in 82% of cases at the 90% level of confidence. However, this clinical limit should be individually set for each BA structures' deviations based on expected impact to target coverage and/or OAR sparing, which is subject to future investigations. Alternatively, the robustness of SPC limits can possibly be improved by also considering the rate of the changes during the treatment to define SPC control limits.

In contrast to the observations for BA deviations, SPC process limits were useful to detect soft tissue trends that exceeded the process limits before the end of treatment week four which could be confirmed in 90% of cases.

DIR precision is important for correct anatomical mapping of the pCT to daily CBCTs. In the absence of a golden standard, trends acquired using DIR were compared with results obtained using an alternative manual method where possible in combination with an uncertainty analyses were conducted. Overall at the 90% level of confidence, BA deviation trends acquired with DIR and MMR overlapped in 77% of cases. This was slightly lower than expected (0.9^2) and may be caused by the assumption that various factors in the uncertainty analyses were normally distributed. Differences in E_0 between DIR and MMR ranged from -1.5 to 1.5 mm (Supplementary Data **A.4:** Individual patient 90% level of confidence SPC results). Further investigation into the differences between the two methods revealed two main causes: 1) Large deviations of mobile structures such as mandible could not always be recovered by DIR; 2) BA structures appeared to have expanded ('creep') after DIR due to structure boundaries not coinciding with anatomical boundaries such as the superior end of the occipital bone,

and due to differences in pCT and CBCT Hounsfield Unit (HU) calibration. For the DIR algorithm used in this study, application of local rigidity constraints to the delineated BA may reduce the impact of 'creep' [197,198].

Precision of PG propagation using DIR was evaluated by comparing these results against independent radiation oncologist CBCT re-contours (Supplementary Data **A.8:** Radiation oncologist PG review) and were found to be in excellent agreement with values reported in literature [138,151,178]. Well aligned to literature [79], the PGs included in this study demonstrated an average volume decrease of 21% and an average medial shift of 2.6 mm during HNRT.

Considering the high correlation between these metrics, there was no indication that V_f^{neck} provides a more accurate metric to assess changes of the treatment volume than overall changes in patient weight.

Dosimetric analysis establishing the clinical relevance of the observed non-rigid changes in pose and anatomy is required to define objective decision rules and appropriate thresholds that can be applied in a safety net for patients who exhibit larger non-rigid deformations in the context of PTV margin reduction. We are therefore currently investigating these aspects in concert with the most efficient treatment adaptation approach for anatomical changes.

In conclusion, this study assessed the potential to quantify and monitor an individual patient's deformation, i.e., non-rigid changes in pose and anatomy during HNRT using DIR in combination with SPC. BA deviation *trends* occurring before the end of treatment week four could only be confirmed in 24% of cases when SPC process limits were used, whereas *absolute* BA deviations could be confirmed in 82% of cases when an *a priori* 2 mm clinical limit was used. SPC process limits were useful to detect soft tissue trends occurring before the end of treatment week four which could be confirmed in 90% of cases. The approach proposed in this study could facilitate timely treatment adaption through detection of problematic patient positioning reproducibility and anatomy changes for individual patients.

3 QUANTIFYING THE DOSE ACCUMULATION UNCERTAINTY AFTER DEFORMABLE IMAGE REGISTRATION IN HEAD-AND-NECK RADIOTHERAPY

Full citation,

- Lowther, N.J., Marsh, S.H. and Louwe, R.J., 2020. Quantifying the dose accumulation uncertainty after deformable image registration in head-and-neck radiotherapy. *Radiotherapy and Oncology, 143,* pp.117-125.

Author contributions,

I developed the intended clinical workflow and the methodology to quantify the dose accumulation uncertainty with the in silico ground truth. I collected all patient data, performed all tasks within the treatment planning system, wrote all the MATLAB scripts for data analysis and performed all tasks within Slicer. I performed all the mathematics for the DVH uncertainty and statistical analyses. I, along with co-authors SM and RL developed the conclusions and refined the methods. I wrote this paper with modifications to structure, flow and readability from SM and RL.

Quantifying the dose accumulation uncertainty after deformable image registration in head-and-neck radiotherapy

Nicholas J. Lowther[1,2], Steven H. Marsh[2], Robert J. W. Louwe[1]

[1]*Wellington Blood and Cancer Centre, Department of Radiation Oncology, Wellington, New Zealand*

[2]*University of Canterbury, School of Physical and Chemical Sciences, Christchurch, New Zealand*

Keywords: dose accumulation uncertainty; dose reconstruction; dose warping; head-and-neck radiotherapy; deformable dose accumulation

Abstract

Background and purpose: Deformable image registration (DIR) facilitated dose reconstruction and accumulation can be applied to assess delivered dose and verify the validity of the treatment plan during treatment. This retrospective study used *in silico* deformations based on clinically observed anatomical changes as ground truth to investigate the uncertainty of reconstructed and accumulated dose in head-and-neck radiotherapy (HNRT).

Materials and methods: A planning CT (*pCT*), cone beam CT (CBCT) from week one of treatment and three later CBCTs were selected for 12 HNRT patients. These images were used to generate *in silico* reference CBCTs and deformation vector fields (DVFs) as ground truth with B-spline DIR. Inverse consistency (IC) of voxels was assessed by determining their net displacement after successive application of the forward and backward DVF. The reconstructed dose based on demons DIR was compared to the ground truth to assess the structure-specific uncertainties of this DIR algorithm for inverse consistent and inverse *in*consistent voxels.

Results:

Overall, 98.5% of voxels were inverse consistent with the 95% level of confidence range for dose reconstruction of a single fraction equal to [-2.3%; +2.1%], [-10.2%; +15.2%] and [-9.5%; +12.5%] relative to their planned dose for target structures, critical organs at risk (OARs) and non-critical OARs, respectively. Inverse *in*consistent voxels generally showed a higher level of uncertainty.

Conclusion:

The uncertainty in accumulated dose using DIR can be accurately quantified and incorporated in dose-volume histograms (DVHs). This method can be used to prospectively assess the adequacy of target coverage during treatment in an objective manner.

3.1 Introduction

The introduction of intensity modulated radiotherapy has enabled highly conformal dose deliveries which allows dose reduction to organs at risk (OAR) and reduced treatment toxicity [21,191,192]. However with highly conformal dose distributions, non-rigid patient deformations (i.e., changes in patient pose, weight loss, tumor response, OAR shift and shrinkage) during treatment [78] may lead to unacceptable loss in target coverage or over-dosage to OARs [93]. Deformable image registration (DIR) facilitates dose reconstruction and subsequent accumulation of the reconstructed dose per fraction to assess total delivered dose to target volumes and OARs in the presence of these anatomical changes [58,107,139,142]. A commonly used dose reconstruction and accumulation procedure matches the planning computed tomography (*pCT*) image (*source*) to an image taken during treatment (*target*) using DIR. This match is described by a 'forward' deformation vector field (DVF). The plan (beam arrangement, monitor units, fluence maps) is then recalculated on the deformed *pCT (dCT)*, and the resulting dose distribution is mapped back to the *pCT* using the inverse or 'backward' DVF. Alternatively, dose reconstruction can be done using Hounsfield unit (HU) corrected on-treatment images, and subsequently map the reconstructed dose to the *pCT* using DIR [199]. Intrinsic limitations of DIR algorithms [134,153] may cause displacement field errors (DFEs) in the 'forward' registration or inverse consistency errors (ICEs) in the 'backward' transformation [159]. These limitations specifically show up in areas of low image contrast and large anatomical changes and result in uncertainties in the reconstructed dose [145,162–165]. The suitability of DIR for accurate and precise dose accumulation is debated in the literature [143,166–169]. A standardized approach to assess the accuracy of DIR facilitated dose accumulation may encourage a large scale implementation.

This investigation includes objectively assessing the adequacy of target coverage using DIR facilitated dose accumulation. This retrospective study was instigated as no literature exists on the accuracy of this specific approach. The objects of this study were to quantify the uncertainty in the reconstructed and accumulated dose [80] based on a DIR algorithm employed by Varian's SmartAdapt [135] specifically, and to provide a method to incorporate these results as a confidence interval in dose-volume histograms (DVHs). The methodology developed in this study utilized *in silico* deformations based on clinically

observed anatomical changes as ground truth to evaluate the dose accumulation uncertainty. This methodology could also be used to validate reconstructed and accumulated doses acquired with other DIR implementations.

3.2 Materials and methods

3.2.1 Patient group

Twelve patients with cancers in the head-and-neck region that were previously treated radically were selected for this retrospective study. The patients' characteristics are summarized in the Supplementary Data (**A.1:** Patient group). Each patient's *pCT* was used to generate a retrospective volumetric modulated arc therapy (VMAT) plan (Eclipse Treatment Planning v11.6 - 13.7, Varian Medical Systems, Palo Alto CA, USA) using 3 mm PTV margin expansions (prescribed dose 54 Gy and simultaneous integrated boost volumes to 60 and 66 Gy in 30 fractions) optimized according to the planning protocol in the Supplementary Data (**A.9:** Planning protocol). The dose calculation grid size was 2.5 x 2.5 x 3.0 mm^3. The *pCT* scans (Brilliance Big Bore; Philips Medical Systems, Eindhoven, The Netherlands) had a voxel size of 1.3 x 1.3 x 3.0 mm^3. Patients were immobilized using a 2.4 mm Reloadable Head and Shoulder S-Frame Kevlar Mask (Q-Fix, Avondale PA, USA) and an individual head and shoulder support vacuum bag (Klarity Medical Products, Newark OH, USA). Daily (n = 30) cone beam computed tomography (CBCT) scans were acquired to verify patient treatment position. The CBCT scans with a voxel size of 0.5 x 0.5 x 2 mm^3 were acquired prior to treatment using a Varian Truebeam (v2.0 or v2.5; Varian Medical Systems, Palo Alto CA, USA). For a minority of treatment fractions, CBCT scans were acquired on a Varian Clinac (v2.1) with a voxel size of 0.7 x 0.7 x 2.5 mm^3. All patients signed an ethical committee provided waiver to use their data in retrospective audits.

3.2.2 Intended clinical workflow

The intended clinical workflow as displayed in Figure 3.1 uses DIR facilitated dose accumulation to objectively assess the appropriateness of the original treatment plan in the context of non-rigid patient deformations. First, a 6 degree of freedom (DOF) rigid bony anatomy registration of the CBCT and the *pCT* was performed. Second, the *pCT* is matched to the pre-treatment CBCT using Varian's demons DIR [133,134] implementation in SmartAdapt (v.13.6, Varian Medical Systems, Palo Alto CA, USA) which produces the *dCT* and is described by a forward DVF. As *dCT*s encompass the correct HU calibration, they can be used to reconstruct the dose delivered during fraction *f* using the beam arrangement, monitor units and fluence maps from the original treatment plan. Third, the original treatment plan is recalculated on the *dCT* and the resulting dose distribution mapped back to the *pCT* space according to the inverse DVF producing the fraction-specific reconstructed dose. Fourth, the reconstructed dose distributions of successive fractions are accumulated. SmartAdapt does not facilitate dose accumulation and the intended workflow requires external software. Therefore, 3D Slicer (v4.8) which is available as freeware was used to accumulate the dose reconstructed for individual fractions [200,201].

Initial assessment of SmartAdapt's forward and inverse (i.e., CBCT-to-*pCT*) registrations highlighted noticeable ICEs especially for large deformations and in areas of low CBCT contrast (Supplementary Data **A.10:** Symmetry and inverse consistency). A consequence of this shortcoming for this evaluation study would be that the resulting dose reconstruction errors would not be assigned to the same voxel as it originated from, which would invalidate the dose accumulation of the reconstructed dose of the individual fractions. It was therefore decided not to use the inverse or symmetric (i.e., interchanging the source and target image) DVF provided by SmartAdapt, but to generate a true inverse DVF to warp the reconstructed dose back to *pCT* space. This true inverse was calculated by inverting the voxel-specific displacement vector described by the Digital Imaging and Communications in Medicine (DICOM) deformable spatial registration module and was also done using 3D Slicer.

3.2.3 Assessing the accuracy of the intended clinical workflow

Figure 3.1b and Figure 3.1c detail the approach that was used to assess the accuracy of the dose reconstruction and accumulation. This approach included an initialization phase (Figure 3.1b) and a dose comparison phase (Figure 3.1c). The dose comparison phase was designed in such a way that it tests whether the clinical workflow can accurately reconstruct the delivered dose for a known reference deformation. These reference deformations were generated *in silico* during the initialization phase and were based on clinically observed deformations. The dose comparison phase included both a test branch (right-hand side; orange) and a reference branch (left-hand side; blue). The test branch is identical to the clinical workflow depicted in Figure 3.1a but uses the known reference deformation with respect to the *pCT*, $dCBCT_f^{ref}$ as input. Exactly the same deformation vector field DVF_f^{ref} that was used in the initialization phase to generate $dCBCT_f^{ref}$ was applied throughout the reference branch of Figure 3.1c. In this way, the reference deformed dose distribution dDD_f^{ref} that was produced by the reference branch could be used as the ground truth against which the deformed dose distribution generated in the test branch dDD_f^{test} was compared.

The initialization phase (Figure 3.1b) included two steps to generate the reference deformation vector field DVF_f^{ref} and $dCBCT_f^{ref}$ of fraction *f* as input for the test branch in the dose comparison phase. The first step of this initialization phase comprised the creation of a virtual CBCT, $CBCT^{plan}$ which closely resembles the *pCT*. This was accomplished by DIR of the *pCT* and a selected CBCT that was acquired during the first week of treatment, $CBCT^{1st\ week}$ using SmartAdapt. In the second step, DVF_f^{ref} was generated by matching $CBCT^{plan}$ to the CBCTs of treatment fractions *f* = 3, 16 and 29 using B-spline DIR. For the current study, the B-spline algorithm was used as the reference DIR algorithm as implemented within the Plastimatch [202] module of 3D Slicer [200,201] using a 2-stage, 2-resolution procedure (stage 1/2: image subsampling rate, 4,4,2/2,2,1; grid size, 100/50 mm; regularization, 0.005; landmark penalty, 0.005; max iterations, 100).

The DVF was limited to the CBCT field of view (FOV) in the inferior-superior direction and is generally not accurate at the last few mm near the inferior-superior CBCT borders. As a consequence, dose

delivered in these regions can either not, or not accurately be mapped back to the *pCT* space. Therefore, *pCT* structures extending outside the superior-inferior extent of the CBCT FOV were cut back 5 mm from the average location of the superior-inferior borders. The in-plane extent of the CBCT FOV was large enough to include all structures in the anterior/posterior and left/right direction for all patients included in this study.

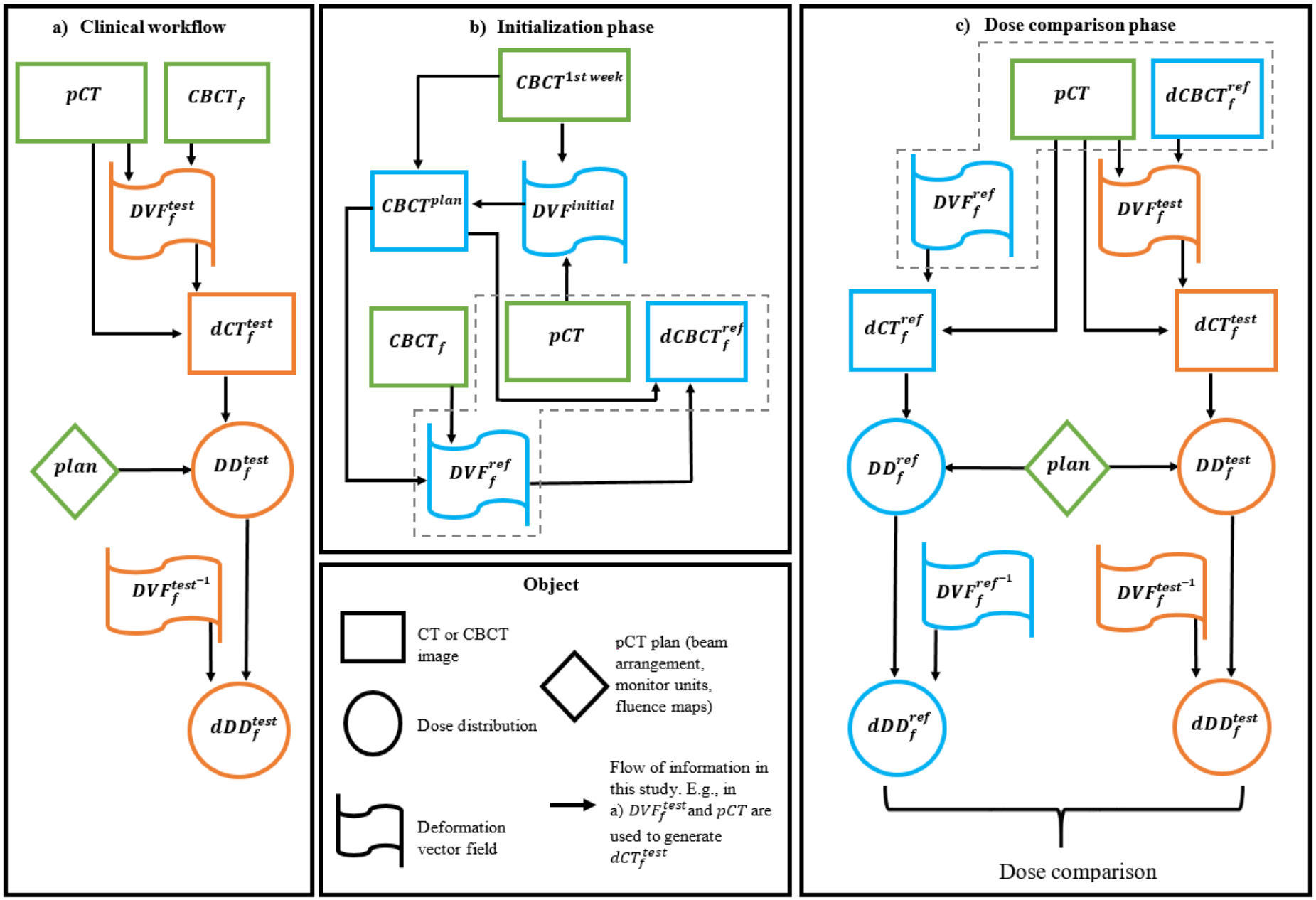

a) Clinical workflow
pCT
CBCT_f
DVF_f^test
dCT_f^test
plan
DD_f^test
DVF_f^test-1
dDD_f^test
b) Initialization phase
CBCT^1st week
CBCT^plan
DVF^initial
CBCT_f
pCT
dCBCT_f^ref
DVF_f^ref
Object
CT or CBCT image
pCT plan (beam arrangement, monitor units, fluence maps)
Dose distribution
Deformation vector field
Flow of information in this study. E.g., in a) DVF_f^test and pCT are used to generate dCT_f^test
c) Dose comparison phase
pCT
dCBCT_f^ref
DVF_f^ref
DVF_f^test
dCT_f^ref
dCT_f^test
DD_f^ref
plan
DD_f^test
DVF_f^ref-1
DVF_f^test-1
dDD_f^ref
dDD_f^test
Dose comparison

Figure 3.1: Framework used to estimate uncertainty in the reconstructed dose as obtained using the demons algorithm. Green, orange and blue outlines represent objects from the clinical workflow, the DIR workflow under investigation, and the reference standard generation, respectively. Panel (a): intended clinical workflow to accumulate dose; (b) initialization phase to generate an input CBCT for the DIR workflow under investigation in this study; (c) dose comparison phase showing the evaluation workflow as conducted in this study. The dashed gray lines in panel (b) and (c) indicate that these represent corresponding process steps.

DIR = deformable image registration. pCT = planning computed tomography. CBCT = cone beam computed tomography dCBCT = deformed CBCT. dCT = deformed pCT. DVF = deformation vector field. DD = dose distribution. dDD = deformed DD. ref = object acquired with B-Spline DIR. test = object acquired with demons DIR in SmartAdapt

3.2.4 Voxel-specific evaluation of observed dose differences

In-house developed software (MATLAB 2018b version 9.5, The MathWorks Inc.) was used to analyze the voxel-specific results of the reconstructed dose distributions dDD_f^{ref} and dDD_f^{test} of the dose comparison phase. First, the uncertainty in the test DIR workflow was assessed by considering the potential error contributions from both branches to the voxel-specific dose difference between the reconstructed dose distributions dDD_f^{ref} and dDD_f^{test}. For that purpose, it was assumed that both branches include the dose reconstruction errors DFE and ICE (Figure 3.2).

$$dDD_f^{test} - dDD_f^{ref} = DFE_f^{test} + ICE_f^{test} - DFE_f^{ref} - ICE_f^{ref} \tag{5}$$

Considering that the known reference deformation vector field DVF_f^{ref} was generated using the reference DIR algorithm, DFE_f^{ref} equals zero per definition. For each head-and-neck radiotherapy (HNRT) structure of interest, the total dose error in the test branch $\left(i.e., DFE_f^{test} + ICE_f^{test}\right)$ was then calculated for each voxel and expressed as a percentage of the planned dose per fraction, Δ^{tot}:

$$\Delta^{tot} = \frac{dDD_f^{test} - dDD_f^{ref} + ICE_f^{ref}}{D_f^{plan}} \times 100\% \tag{6}$$

The magnitude of ICE_f^{ref} was assessed by applying DVF_f^{ref} to dDD_f^{ref}, and subtracting the result from DD_f^{ref}.

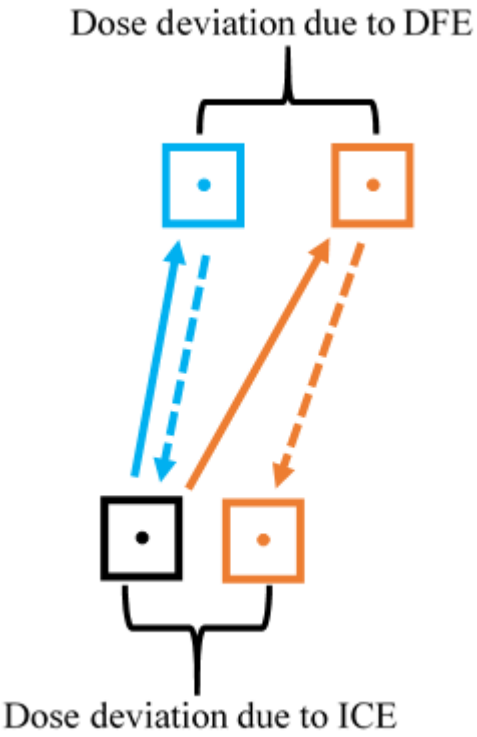


Figure 3.2: Schematic representation of the factors contributing to the total dose difference of a specific voxel (black square) between the test (orange pathway) and reference (light blue pathway) deformable image registration (DIR) dose reconstructions. The 'forward' registrations are shown by the solid arrows. The displacement field error (DFE) of the test registration (orange pathway) is exemplified by the top light blue and orange voxels not perfectly aligning. The 'backward' registration is shown by the dashed arrow. The inverse consistency error (ICE) of the test registration is exemplified by the bottom black and orange voxels not perfectly overlapping.

3.2.5 Assessing DVH uncertainty in the clinical workflow based on the results of this study

After each voxel's contribution to the reconstructed test dose uncertainty was determined as described above, each voxel was stratified by whether its forward displacement vector was inversely consistent (referred to as 'inverse consistent voxels' from this point onward) or not ('inverse *in*consistent voxels'). The methodology applied to make this distinction can also be applied in a clinical setting, and allows summation of the voxel-specific uncertainty over multiple fractions using the estimated uncertainty from this study based on this distinction. The displacement error $dICE_f^{test}$ was calculated by subsequently applying DVF_f^{test} and the inverse DVF_f^{test} provided by SmartAdapt to the *pCT*, and subtracting this result from the original *pCT*. Subsequently, an inverse consistency (IC) threshold equal to the size of one voxel of the dose calculation grid (i.e., 2.5 x 2.5 x 3.0 mm^3), V_{sz}, was used to distinguish between inverse consistent and inverse *in*consistent voxels of $dICE_f^{test}$. The IC rate was defined as the number of voxels meeting the IC threshold criterion divided by the total number of voxels, and was used to assess the validity of the dose reconstruction uncertainty estimates in this study. In a clinical setting, the IC rate could potentially be used as additional information to assess the validity of the reconstructed DVH. Considering that the dose distributions and dose gradients of various structures in HNRT are generally distinct, the dose reconstruction uncertainties were evaluated separately for each structure type.

For each structure *S*, the dose reconstruction uncertainties of inverse consistent u_c and inverse *in*consistent u_i voxels were estimated using the 95% percentile range of the distributions of Δ^{tot} in accordance with the Guide to the Expression of Uncertainty in Measurement (GUM) [196]. In addition, the upper and lower part of the 95% range were assessed separately to maintain the skewed, non-normal distributed character of Δ^{tot}. For dose reconstruction in the intended clinical workflow, the 95% confidence interval of the accumulated dDD_f^{test} dose for a single voxel within structure *S* at fraction *f*, uA_f^S was calculated according to the following.

Uncertainty summation in quadrature assuming a fraction-to-fraction normal distribution of *u*:

$$uA_f\ [Gy] = \sqrt{\sum_{i=1}^{f}(D_i^r\ .\ u)^2} \tag{7a}$$

For a system with two uncertainty levels this can be expanded into:

$$uA_f\ [Gy] = \sqrt{\sum_{i=1}^{f}(D_i^r\ .\ u)^2} \qquad u = \begin{cases} u_c & \text{if } dICE_f^{test} < V_{sz} \\ u_i & \text{if } dICE_f^{test} > V_{sz} \end{cases} \tag{7b}$$

Expressed as a percentage of accumulated dose for separate upper and lower levels of uncertainty:

$$uA_f^{S+}\ [\%] = \frac{\sqrt{\sum_{i=1}^{f}(D_i^r\ .\ u)^2}}{D_f^a} \times 100\% \qquad u = \begin{cases} u_c^{S+} & \text{if } dICE_f^{test} < V_{sz} \\ u_i^{S+} & \text{if } dICE_f^{test} > V_{sz} \end{cases} \tag{7c}$$

$$uA_f^{S-}\ [\%] = \frac{\sqrt{\sum_{i=1}^{f}(D_i^r\ .\ u)^2}}{D_f^a} \times 100\% \qquad u = \begin{cases} u_c^{S-} & \text{if } dICE_f^{test} < V_{sz} \\ u_i^{S-} & \text{if } dICE_f^{test} > V_{sz} \end{cases} \tag{7d}$$

where D_i^r are the reconstructed doses of each fraction and D_f^a is the accumulated dose for fraction *f*. The 95% level of confidence for each dose level bin in the (cumulative) DVH of the accumulated dose was calculated by averaging the uA_f of all voxels with D_f^a equal to or larger than that dose level. As an example of incorporating the uncertainty in the accumulated dose as a 95% level of confidence interval in DVHs during treatment, the accumulated DVHs including the 95% confidence interval at fraction 30 was calculated for the high-dose PTV, the high-dose clinical target volume (CTV) and the ipsilateral parotid gland (PG) as per the intended dose accumulation clinical workflow (Figure 3.1a) for patient 1.

3.3 Results

3.3.1 Voxel-specific evaluation

Table 5 summarizes the averages and ranges of the IC rate for each of the 14 structures included in the analysis. In addition, Table 5 summarizes the averages and 95% percentile ranges of Δ^{tot} for both inverse consistent and inverse *in*consistent voxels for all structures. The average IC rate for each structure was generally larger than 99% except for the brainstem planning risk volume (PRV), spinal cord PRV, brainstem, spinal cord and ipsilateral PG where the average IC rate was slightly lower but still larger than 94%. More detail about the IC rates for all patients' structures can be found in the Supplementary Data (**A.11:** IC rates).

The distributions of Δ^{tot} values for inverse consistent and inverse *in*consistent voxels is shown in the histograms of Figure 3.3a and Figure 3.3b, respectively. Q-Q probability plots showed that Δ^{tot} was generally not normally distributed (Supplementary Data **A.12:** Q-Q probability plots). Therefore, the standard uncertainty (i.e., u_c and u_i) of the reconstructed dose was estimated using the 95% percentile range of these distributions rather than 1 standard deviation. For inverse consistent voxels, the median Δ^{tot} over all structures was 0.0% with a 95% percentile range equal to (-3.6%; +4.6%). For inverse *in*consistent voxels, the median Δ^{tot} was -0.1% with a 95% percentile range equal to (-10.7%; +10.5%).

For individual structure types, the largest 95% percentile ranges of Δ^{tot} for both inverse consistent and inverse *in*consistent voxels were observed for the brainstem (and brainstem PRV) and PGs. The 95% percentile ranges of Δ^{tot} for inverse consistent and inverse *in*consistent voxels of the brainstem were (-15.6%; +21.0%) and (-15.8%; +14.2%), respectively. For the ipsilateral PG, the 95% percentile range of Δ^{tot} for inverse consistent and inverse *in*consistent voxels was (-10.5%; +14.4%) and (-18.8%; +8.1%), respectively. For the contralateral PG, these were (-10.4%; +12.7%) and (-23.5%; +4.5%), respectively. For all target structures (i.e., PTVs and CTVs), the largest 95% percentile ranges of Δ^{tot} were (-3.1%; +2.7%) and (-6.1%; +3.0%) for inverse consistent and inverse *in*consistent voxels, respectively. Individual structure distributions of Δ^{tot} are available in the Supplementary Data (**A.13:** Histograms of Δ^{tot}).

Table 5: Structure-specific average and range for the inverse consistency (IC) rate, and median and 95% percentile range of Δ^{tot} for inverse consistent and inverse inconsistent voxels. PTV = planning target volume. CTV = clinical target volume. PRV = planning risk volume. PG = parotid gland. SMG = submandibular gland.

		IC rate (%)		Inverse consistent voxels Δ^{tot} (%)		Inverse *in*consistent voxels Δ^{tot} (%)	
Structure	**n**	**mean**	**range**	**median**	**95% percentile range**	**median**	**95% percentile range**
High-dose PTV	11	99.1	85.6 – 100.0	0.0	-2.0; +1.6	0.2	-2.8; +7.7
Intermediate-dose PTV	4	100.0	99.9 – 100.0	-0.1	-2.9; +1.4	0.2	-1.4; +1.4
Low-dose PTV	12	99.5	94.0 – 100.0	0.0	-3.1; +2.7	-0.4	-6.1; +3.0
High-dose CTV	11	99.1	85.3 – 100.0	0.1	-1.3; +1.5	0.2	-2.2; +4.0
Intermediate-dose CTV	4	99.9	99.8 – 100.0	0.0	-1.7; +1.2	0.1	-1.5; +1.4
Low-dose CTV	12	99.4	91.2 – 100.0	0.0	-1.9; +2.0	-0.2	-4.5; +3.3
Brainstem PRV	12	96.9	73.3 – 100.0	0.8	-13.7; +19.5	-0.3	-15.8; +24.6
Spinal cord PRV	12	97.9	78.9 – 100.0	0.4	-6.7; +8.5	1.0	-8.4; +14.9
Brainstem	12	94.1	64.2 – 100.0	1.0	-15.6; +21.0	0.0	-15.8; +14.2
Spinal cord	12	97.1	71.9 – 100.0	0.4	-7.2; +9.0	0.9	-8.1; +15.8
Ipsilateral PG	14	96.8	26.3 – 100.0	0.6	-10.5; +14.4	-1.7	-18.8; +8.1
Contralateral PG	10	99.2	86.2 – 100.0	0.5	-10.4; +12.7	-6.8	-23.5; +4.5
Ipsilateral SMG	13	99.9	95.9 – 100.0	-0.1	-5.5; +3.3	0.7	-0.9; +1.3
Contralateral SMG	10	99.9	97.6 – 100.0	-0.1	-3.6; +5.5	-0.7	-6.4; +0.5
All structures	149	98.5	26.3 – 100.0	0.0	-3.6; +4.6	-0.1	-10.7; +10.5

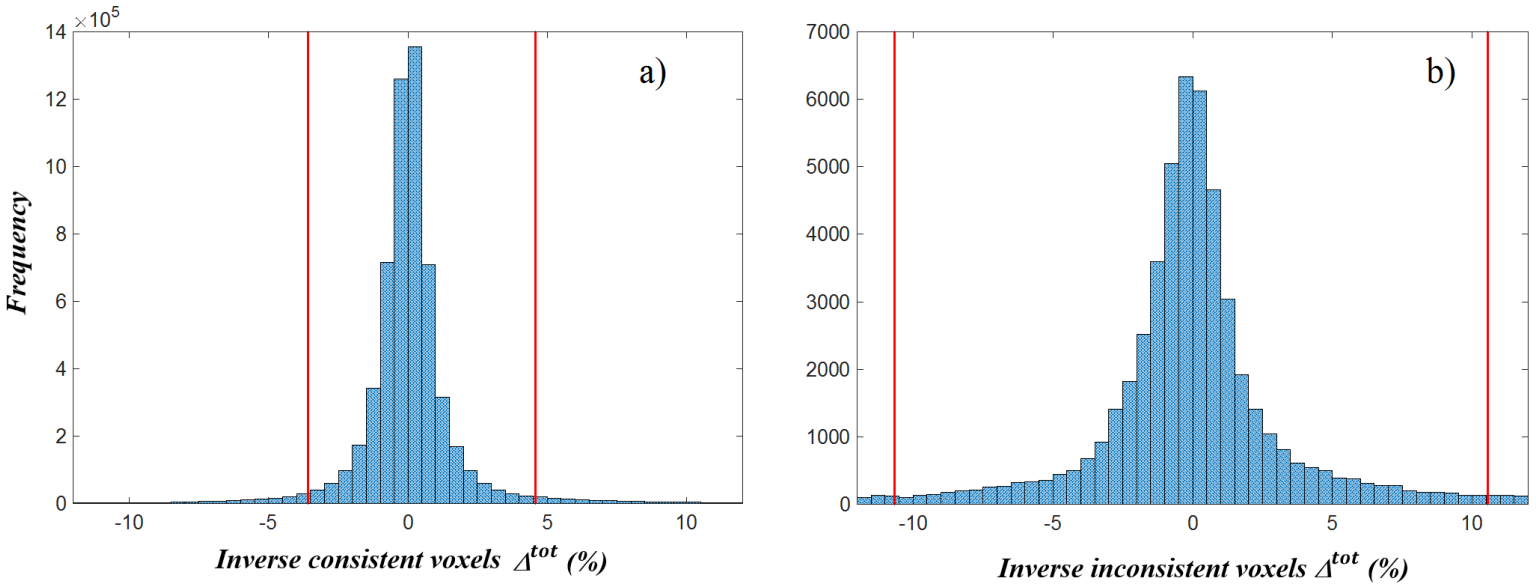


Figure 3.3: Distribution of Δ^{tot} values for inverse consistent (a) and inverse inconsistent (b) voxels from all structures and all patients.

3.3.2 DVH uncertainty example

Over all patients, uA_{30} was (-0.4%; +0.3%), (-0.2%; +0.3%) and (-1.9%; +2.6%) for the inverse consistent voxels within the high-dose PTV, the high-dose CTV and ipsilateral PG, respectively. Note that as per Equation 7 (page 108), these structures' ranges are approximately a factor five smaller than the ranges listed in Table 5. For inverse *in*consistent voxels, uA_{30} for these structures was (-0.5%; +1.4%), (-0.4%; +0.7%) and (-3.4%; +1.5%), respectively.

For patient 1, the DVH metrics for the accumulated dose of the high-dose target volumes were significantly different at the 95% level of confidence compared to the planned values (Figure 3.4a). $D_{98\%}$ (PTV) and $D_{99\%}$ (CTV) were 97.8% and 99.5% of the planned values, respectively. The $D_{50\%}$ values for the high-dose PTV and CTV accumulated dose values were both 100.4% of their planned $D_{50\%}$. The accumulated dose to the ipsilateral PG was only marginally higher than the planned dose (Figure 3.4b).

Figure 3.5 shows the spatial distribution of the reconstructed dose uncertainty (i.e., Δ^{tot}) at fractions 3 and 29 for patient 1's high-dose PTV and ipsilateral PG calculated over all voxels (left column), and for inverse consistent voxels only (right column). Typically, a high IC rate is observed for most structures as can be observed by comparing the left-hand and right-hand columns of Figure 3.5a-f. In addition, a low level of uncertainty in reconstructed dose is typically observed as highlighted in Figure 3.5a-b and Figure 3.5e-f. In a small number of cases an uncertainty larger than 10% in the reconstructed dose was observed, as indicated by the region with the red color wash in Figure 3.5c and Figure 3.5d. The higher uncertainties in these cases were caused by DFEs of the test algorithm in high dose gradient regions and occurred independent of IC. Considerable reduction in the external contour and/or poor CBCT image contrast challenged the IC functionality of the test DIR algorithm and resulted in an exceptionally low IC rate of 26.3% for the ipsilateral PG at fraction 29 (Figure 3.5h). Further investigation found this large contour reduction between the pCT and $CBCT_{29}$ also challenged the reference algorithm, resulting in a non-realistic deformation DVF_{29}^{ref} in the ipsilateral PG region. The test DIR was unable to reproduce this unrealistic reference deformation in an inverse consistent manner. The combination of a low IC rate, high dose gradient and notable DFEs resulted in distinctively higher uncertainty of the reconstructed structure dose for this fraction (Figure 3.5g).

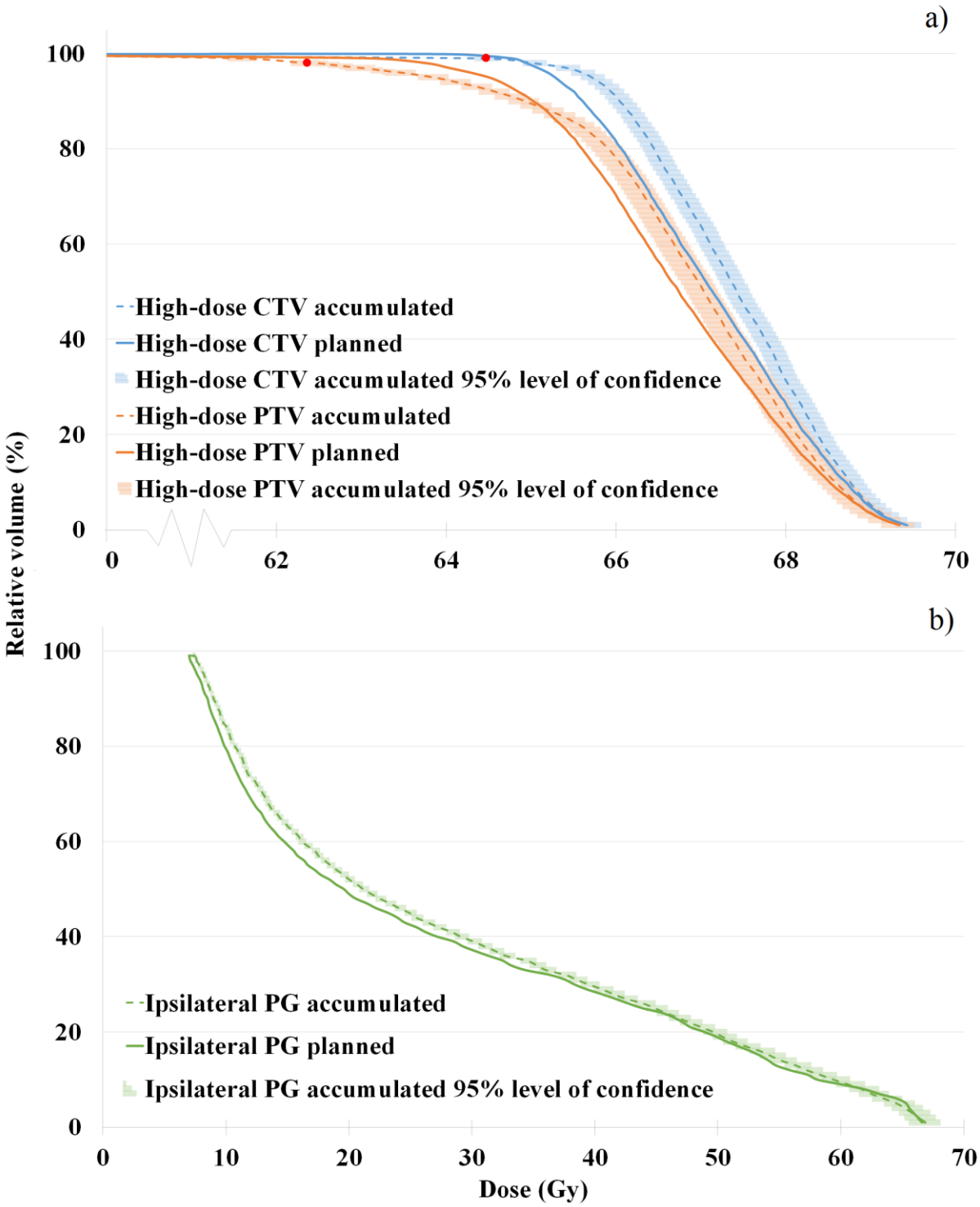


Figure 3.4: Dose-volume histogram (DVH) of the accumulated dose for patient 1's high-dose CTV and PTV (a) and ipsilateral parotid gland (PG) (b) at the end of the 30 fraction treatment compared to the original planning DVHs. The accumulated $D_{98\%}$ and $D_{99\%}$ (red bullet points) of the high-dose PTV and CTV were significantly different at the 95% level of confidence from the planned values but still 97.8% and 99.5% of the planned $D_{98\%}$ and $D_{99\%}$, respectively.

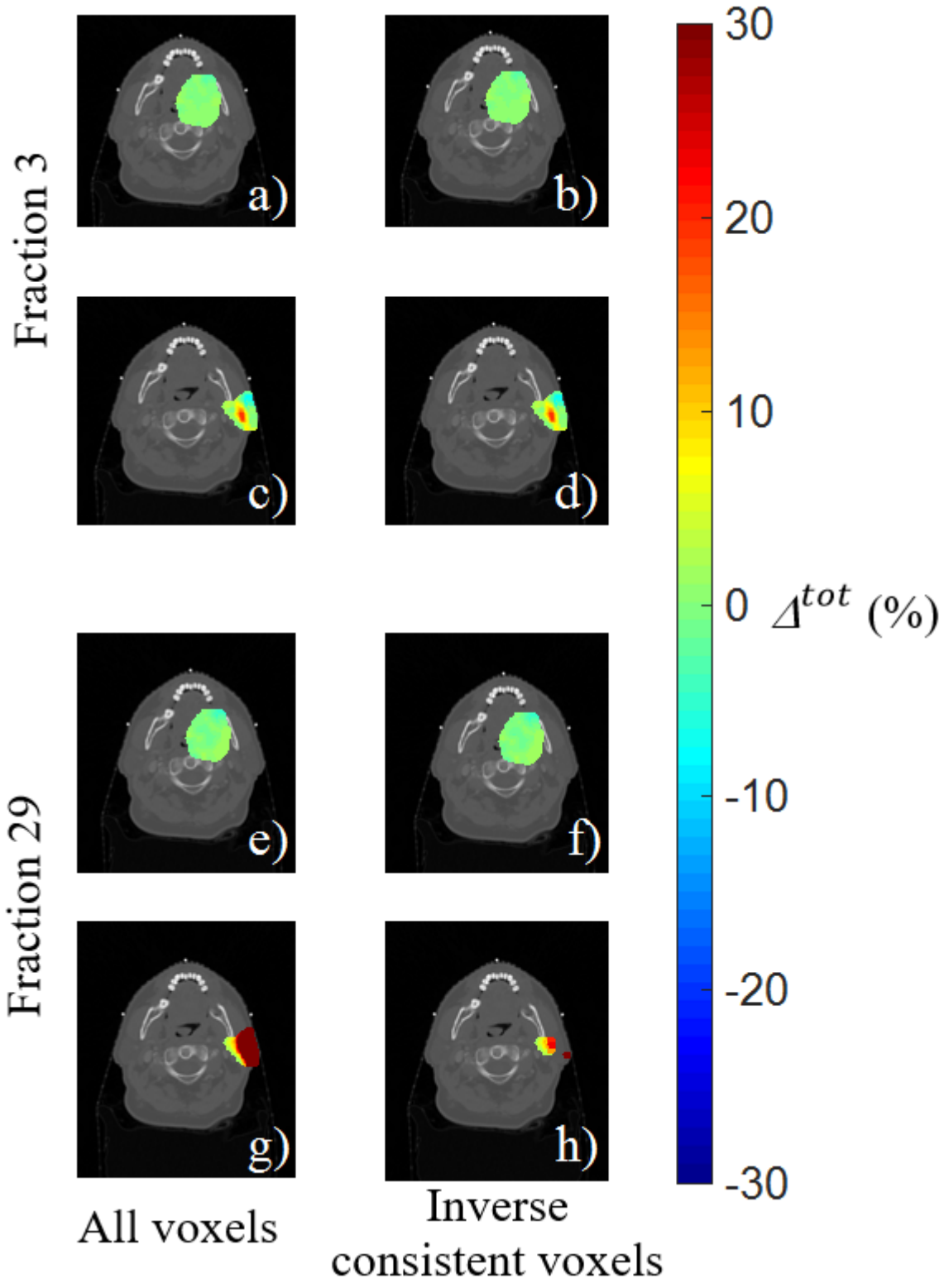


Figure 3.5: Δ^{tot} of all voxels (left column) and inverse consistent voxels only (right column) for patient 1 at fraction 3 (a-d) and fraction 29 (e-h). The high-dose PTV (a,b,e,f) and ipsilateral parotid gland (c,d,g,h) are shown for a single axial slice.

3.4 Discussion

Accurate DIR facilitated dose accumulation for individual patients would be extremely useful to verify the adequacy of the delivered dose in the presence of non-rigid patient deformations (i.e., changes in patient pose, weight loss, tumor response, OAR shift and shrinkage) during treatment [139], and could provide an objective tool to decide whether treatment adaptation is required. In particular when a PTV margin reduction is implemented [1,3,4], loss in target coverage during treatment might occur more often and might become more relevant compared to an approach using standard PTV margins [110,142]. Furthermore, the application of DIR facilitated dose accumulation to monitor individual patients' actual delivered dose during treatment can largely be automated and is therefore attractive from an economic and efficiency perspective. The current study was carried out for an image-guided radiotherapy (IGRT) workflow using the DIR algorithm available at our institute and was based on daily CBCT, but the approach can also be used for other workflows. For instance, the generation of an *in silico* ground truth as demonstrated here, including the procedure to generate a known CBCT deformation (Figure 3.1b) can also be used to validate other DIR algorithms. Alternatively, this procedure can be amended to validate DIR dose accumulation pathways that utilize other imaging schemes (i.e., not CT-to-CBCT).

Limitations of the DIR facilitated dose reconstruction include reduced accuracy due to low image contrast and potential algorithm restrictions to recover large deformations [135]. In addition, DIR algorithms do generally assume that mass is conserved and are therefore not able to correctly replicate volume changes. In particular in high dose gradient areas, low image contrast and large anatomical changes may cause low IC rates and large DFEs in the DIR facilitated dose reconstruction and accumulation approach [139]. This study investigated the uncertainty in DIR facilitated dose reconstruction and accumulation, using *in silico* deformations based on clinically observed anatomical changes. By quantifying the observed dose reconstruction uncertainties for specific structures, and for inverse consistent and inverse *in*consistent voxels separately, this data can be used in a prospective way. This application enables assessing the impact of anatomical changes to target coverage and/or OAR sparing at the 95% level of confidence during treatment.

Many studies assessed the delivered dose based on dose reconstruction using on-treatment imaging [142,199,203]. The accuracy of the dose reconstruction is often estimated based on the geometrical

accuracy of DIR [142,199], while only a few papers directly assessed the dose reconstruction accuracy [145,162,164,204]. To the best of our knowledge, there are no other studies that investigated a demons DIR pCT-CBCT workflow to assess the delivered dose during HNRT, included an uncertainty estimate based on clinically observed deformations as ground truth, and presented a methodology that can be used in a prospective way (i.e., included in on-treatment DVHs).

Using an IC threshold of one dose calculation voxel size, the IC rate was at least 95% for the majority of structures in this study. Only 8% of the 149 structures that were each analyzed for three fractions in this study had an IC rate less than 95% (Supplementary Data **A.11:** IC rates), with the majority of these cases observed for one specific patient. In particular the large external contour reduction of fraction 29 for the ipsilateral PG of this patient as illustrated in Figure 3.5h likely resulted in the extremely low IC rate of 26.3%. The main purpose of the current study was to provide a robust estimate of the dose reconstruction uncertainty by separately assessing the uncertainty for inverse consistent and inverse *in*consistent voxels for various structure types. However, for other DIR algorithms and/or other body sites, the dose reconstruction uncertainty for inverse *in*consistent voxels may be much larger or more variable. In that case an alternative approach could be to use the IC rate in combination with an acceptance threshold for the IC rate that depends on the clinical risk associated with accepting an erroneous dose reconstruction. For implementation of this methodology in other departments, it is up to the discretion of the users to select an approach that works optimal for a specific application.

Larger DFEs were specifically observed for brainstem and the PGs. Further investigation (data not shown) did not reveal a statistically significant trend of the DFE with the progress of treatment for the ipsilateral or contralateral PGs. It was therefore concluded that low image contrast in the CBCTs was a major contributor to the larger DFEs for these structures rather than contour or other anatomical changes. In spite of the lower IC rates and consequently higher estimated dose reconstruction uncertainty in these cases, the example DVHs in Figure 3.4 indicate that the accuracy of dose reconstruction appears to be sufficient for clinical application. Furthermore, it is anticipated that future improvements of the CBCT image quality such as the implementation of iterative reconstruction [205] will not only be beneficial for IGRT in general but may also increase the accuracy of dose

reconstruction. Although it should be noted that the study by Veiga et al. found that CBCT image quality was sufficient for dose of the day calculations based on CT-to-CBCT DIR [206].

Both the reference DIR algorithm and the test DIR algorithm under investigation in this study do not have provisions to account for either the loss or increase of mass of structures. This is a limitation for uncertainty estimates in this study although the two algorithms likely respond differently to volume changes. Furthermore, volume changes will generally reduce the IC rate and result in a higher estimated dose reconstruction uncertainty in prospective applications. Thus, the impact of volume changes on the dose reconstruction accuracy is at least partially included in the uncertainty estimate. A previous study from our group [207] revealed that seven of the patients included in the current study showed a significant change in neck contour over the treatment course which could potentially have impacted the uncertainty estimate for the elective dose target volumes. Similarly, five patients showed at least a 25% decrease of PG volume but volume changes of the GTVs could not be determined due to CBCT image quality restrictions. Literature on the impact of volume changes on the accuracy of DIR based dose reconstruction is scarce but a short debate on this subject concluded that the impact was expected to be limited [167,168]. Future studies are required to quantitatively assess the impact on patient dosimetry of this limitation of the current DIR algorithms but would at least require high contrast imaging during treatment.

The distinction between inverse consistent and inverse *in*consistent voxels to estimate the uncertainty of the accumulated dose in the intended clinical workflow was based on the forward and inverse DVF as provided by SmartAdapt. However, the true inverse DVF was applied in the intended clinical workflow of Figure 3.1a to warp the reconstructed dose. The contribution to the uncertainty of the accumulated dose of inverse inconsistent voxels in prospective application of DIR based dose reconstruction in the intended clinical workflow will therefore be conservative.

It should be noted that the degree to which the reference deformations of the *in silico* ground truth are representative is limited by the accuracy of the reference registrations (i.e., DVF_f^{ref}) which in theory could reduce the applicability of the dose accumulation uncertainties derived here, and alternative DIR algorithms may be able to generate more representative DVFs [203,208]. However, perfectly accurate ground truth deformations are not strictly necessary assuming they are plausible with anatomical

changes routinely seen during treatment and that the reconstructed dose of a test DIR algorithm (e.g., dDD_f^{test}) is evaluated against the reconstructed dose of the reference DIR algorithm (e.g., dDD_f^{ref}). Degraded CBCT image quality, erroneous initial rigid alignment, large initial deformations, etc. could challenge the intrinsic accuracy of DIR algorithms which may lead to systematic residual deformable registration errors. If this occurs at the stages where the reference (ground truth) deformation is generated (Figure 3.1b), the impact of these phenomena on the clinical DIR workflow under investigation would not be fully assessed and could potentially create a bias in results. An extensive visual comparison of *pCT* and $CBCT^{plan}$ assess any systematic residual deformations showed that the CBCT image quality was generally degraded around dental implants due to differences in scatter, which obscured the comparison at some locations. Although no residual deformation errors were found, it cannot be excluded that the dose reconstruction error of the DIR algorithm under investigation around dental fillings is larger than estimated in the current study. In addition, a difference in tongue position between *pCT* and $CBCT^{plan}$ was observed due to swallowing, but this was for one patient only and it is therefore unlikely that this created a bias in results. Further investigation using a considerably less representative CBCT of the first week of treatment as 'worst' case (i.e., notable swallowing differences relative to the *pCT*) to generate an alternative $CBCT^{plan}$ resulted in very comparable 95% percentile ranges of Δ^{tot}. Therefore, the impact of residual differences between the *pCT* and $CBCT^{plan}$ on the estimated dose reconstruction uncertainty are estimated to be negligible. Comparison of $dCBCT_f^{ref}$ and $CBCT_f$ did reveal some residual deviations at a number of locations but these were generally smaller than a dose calculation voxel. However, these residual deformations were not systematic in nature, and it is therefore estimated that a bias in results due to DIR limitations in this part of the workflow is negligible.

The dose difference Δ^{tot} was not available for those voxels which were cut from structures extending outside the superior-inferior extent of the CBCT due to the limited CBCT FOV. However, it is reasonable to assume that the measured structure-specific 95% percentile range of Δ^{tot} is representative for the actual 95% percentile range considering that it is unlikely that residual deformations and dose reconstruction errors will be different outside the CBCT FOV, and that the Δ^{tot} values were assessed over multiple patients.

Two studies [145,162] have evaluated uncertainties in the dose reconstruction and accumulation process for HNRT specifically. The study of Veiga et al. [145] compared these uncertainties for four DIR algorithms and five HNRT patients based on CT-to-CBCT registration. These DIR algorithms represented three different implementations of B-spline which were compared against the results obtained using a symmetric B-spline registration parameterized by a stationary velocity field. The latter algorithm was chosen as the reference algorithm for that study due to its inverse and symmetry consistency which are properties theoretically associated with physically plausible patient deformations during treatment [130,131,146]. However, no ground truth was defined which may have confounded the reconstructed dose comparison. The best performing DIR algorithm in this study recorded a root mean square (RMS) dose difference of 1.6% and a 95% percentile range of ±1.4% with respect to the prescribed dose for all voxels encompassed by the 95% isodose surface. This is very comparable to the results of our study where the median u_c for the high-dose PTV was 0.0% with a 95% percentile range equal to (-2.0%; +1.6%).

The study of Rigaud et al. [162] compared ten DIR approaches to register the planning CT with weekly repeat CTs for 15 HNRT patients. The uncertainty in dose accumulation was then assessed based on the registration errors of 14 landmarks as identified by an expert. Of the 10 DIR approaches in this study, the B-spline DIR algorithm utilizing mutual information on filtered CTs performed best with an average dose difference error of 1.7% and a 95% percentile range of ±3.8% relative to the planned dose, respectively. Again, this compares well with the results of our study where the overall median u_c was 0.0% with a 95% percentile range equal to (-3.6%; +4.6%).

This study assessed the structure-specific uncertainties in DIR based dose reconstruction during HNRT using *in silico* deformations based on clinically observed anatomical changes as ground truth. This uncertainty was quantified both for inverse consistent and inverse *in*consistent voxels, as well as the average IC rate for each structure. This data can be used in a prospective way to estimate the accuracy of DIR based dose reconstruction and provides a method to objectively assess the adequacy of treatment in the context of anatomical changes or variations in patient positioning. The results of this study suggest that DIR based dose accumulation is sufficiently accurate to assess the adequacy of target coverage during HNRT.

4 DOSE ACCUMULATION TO ASSESS THE VALIDITY OF TREATMENT PLANS WITH REDUCED MARGINS IN RADIOTHERAPY OF HEAD AND NECK CANCER

Full citation,

- Lowther, N.J., Marsh, S.H. and Louwe, R.J., 2020. Dose accumulation to assess the validity of treatment plans with reduced margins in radiotherapy of head and neck cancer. *In publication in Physics and Imaging in Radiation Oncology*

Author contributions,

I applied the methodology that I developed in *Chapter Three* to simulate the intended clinical workflow for HNRT at the WBCC. I collected all patient data, performed all tasks within the treatment planning system, wrote all the MATLAB scripts for data analysis and performed all tasks within Slicer. I performed all the mathematics for the statistical analyses. I, along with co-authors SM and RL developed the conclusions and refined the methods. I wrote this paper with modifications to structure, flow and readability from SM and RL.

Dose accumulation to assess the validity of treatment plans with reduced margins in radiotherapy of head and neck cancer

Nicholas J. Lowther[1,2], Steven H. Marsh[2], Robert J. W. Louwe[1*]

[1]*Wellington Blood and Cancer Centre, Department of Radiation Oncology, Wellington, New Zealand*

[2]*University of Canterbury, School of Physical and Chemical Sciences, Christchurch, New Zealand*

Keywords: head-and-neck radiotherapy; DIR dose accumulation; PTV margin reduction; plan robustness; anatomical changes

Abstract

Background and purpose: Literature has reported reduced treatment toxicity in head-and-neck radiotherapy (HNRT) when reducing the planning target volume (PTV) margin from 5 to 3 mm but loco-regional control was not always preserved. This study used deformable image registration (DIR)-facilitated dose accumulation to assess clinical target volume (CTV) coverage in the presence of anatomical changes.

Materials and methods: VMAT plans for 12 patients were optimized using 3 or 5 mm PTV and planning risk volume (PRV) margins. The planning computed tomography (*pCT*) scan was registered to each daily cone beam CT (CBCT) using DIR. The inverse registration was used to reconstruct and accumulate dose (D^{acc}). CTV coverage was assessed using the dose-volume histogram (DVH) metric $D_{99\%}^{acc}$ and by individual voxel analysis. Both approaches included an uncertainty estimate using the 95% level of confidence.

Results: $D_{99\%}^{acc}$ was less than 95% of the prescribed dose D^{presc} for three cases including only one case where this was at the 95% level of confidence. However for many patients, the accumulated dose included a substantial volume of voxels receiving less than 95% D^{presc} independent of margin expansion, which predominantly occurred in the subdermal region. Loss in target coverage was very patient specific but tightness of target volume coverage at planning was a common factor leading to underdosage.

Conclusion: This study agrees with previous literature that PTV/PRV margin reduction did not significantly reduce CTV coverage during treatment, but also highlighted that tight coverage of target volumes at planning increases the risk of clinically unacceptable dose delivery. Patient specific verification of dose delivery assessing the dose delivered to each voxel is recommended.

4.1 Introduction

The introduction of intensity modulated radiotherapy has enabled highly conformal dose deliveries which allow dose reduction to organs at risk (OARs), and result in reduced treatment toxicity [21,191,192]. These highly conformal treatments require image-guided radiotherapy (IGRT) to warrant accurate patient positioning and monitoring of changes in patient anatomy [86]. Planning target volume (PTV) margins are applied to target volumes [170] during treatment planning to account for uncertainties such as patient positioning, geometrical accuracy of the treatment machine and geometrical uncertainties of target volume definition. The presence of these uncertainties have also prompted a recommendation to apply a planning risk volume (PRV) margin to critical OARs [42]. PTV and PRV margins of 5 mm are commonly applied in head-and-neck radiotherapy (HNRT) [171,209,210]. However, PTV margin recipes [35,36] do not account for non-rigid anatomy changes (i.e., changes in patient pose, weight loss, tumor response, OAR shift and shrinkage) which are commonly observed in HNRT [78,79,207]. Nevertheless, two groups have reported favorable toxicity profiles while maintaining good two-year loco-regional (LR) control rates after reducing the PTV margin from 5 to 3 mm [1–3]. More recently, a third study by Franzese et al. [4] reported a significant difference in the two-year LR control rates between patients treated with 5 or 3 mm PTV margins. In the latter study, a 3 mm PTV margin was associated with a decreased rate of LR control. A retrospective study by van Kranen et al. [142] found a slight increase in the risk of clinical target volume (CTV) underdosage when reducing the PTV margin from 5 to 3 mm while a similar study by Wu et al [110] concluded that the coverage of the CTVs at the end of treatment was not affected by a PTV margin reduction. Considering these differing results, further investigation of risk factors that could jeopardize a patient's CTV coverage and/or critical OAR avoidance in a reduced margin setting is warranted.

This study utilized deformable image registration (DIR) facilitated dose accumulation to objectively assess the adequacy of target coverage, critical OAR avoidance and non-critical OAR sparing when either 5 or 3 mm PTV and PRV margins were applied. In addition, the occurrence, location and trends where CTV coverage and critical OAR avoidance could be at risk was investigated. This study contributes to the scarce body of literature regarding the robustness of treatment plans for anatomical

changes during HNRT when margins are reduced [110,142]. While used retrospectively in this study, the presented dose accumulation analysis is primarily intended for prospective clinical application.

4.2 Materials and methods

4.2.1 Treatment immobilization, planning and on-treatment imaging

Twelve patients with cancers in the head-and-neck region that were previously treated radically were selected for this study. The patients' characteristics are summarized in Supplementary Data **A.1:** Patient group. The planning computed tomography (*pCT*) scans (Brilliance Big Bore; Philips Medical Systems, Eindhoven, The Netherlands) had a voxel size of 1.3 x 1.3 x 3.0 mm^3. Patients were immobilized using a 2.4 mm Reloadable Head and Shoulder S-Frame Kevlar Mask (Q-Fix, Avondale PA, USA) and an individual head and shoulder support vacuum bag (Klarity Medical Products, Newark OH, USA).

Each patient's *pCT* was used to generate two volumetric modulated arc therapy (VMAT) plans (Eclipse Treatment Planning, Varian Medical Systems, Palo Alto CA, USA) using either 5 or 3 mm PTV and PRV margin expansions from the CTVs and critical OARs (i.e., brainstem and spinal cord), respectively. Plan optimization was carried out using the protocol in Supplementary Data **A.9:** Planning protocol, using a dose calculation grid size of 2.5 x 2.5 x 3.0 mm^3. For eleven patients, a 5 mm expansion of the gross tumor volume (GTV) was used to generate the high-risk CTV with a prescribed dose of 66 Gy in all but one case where a compartmentalization approach [34] was used. For one patient with a benign tumor, the prescribed dose to the target volume was 54 Gy and the CTV was created by a 3 mm expansion of the GTV (see Supplementary Data **A.1:** Patient group and **A.14:** Structures available for analysis). For four patients, an intermediate-risk CTV was defined with a prescribed dose of 60 Gy to include anatomical structures with a high probability of infiltration based on positron-emission tomography (PET) imaging and experience of the radiation oncologist. The low-risk CTV included structures as per The Danish Head and Neck Cancer Group (DAHANCA) consensus guidelines [176] and was planned to 54 Gy. All plans used two full arcs to create a highly conformal dose distribution around the 54 Gy low-risk CTV and simultaneous integrated boost volumes to 60 and 66 Gy in 30 fractions. To achieve similar plans in terms of conformity and avoid a bias in the plan comparison, the first author (NL) generated each patient's 3 and 5 mm plan in immediate succession. All plans fulfilled the departmental criteria for treatment plan acceptance which are based on the studies of Doornaert et al. [47] and Verbakel et al. [29] and adhere to ICRU guidelines [211]. Plans were reviewed by a senior clinician involved in HNRT and a senior medical physicist who deemed the plans clinically acceptable.

In a small number of cases where a CTV was located superficially within 3 mm of the external contour without the use of bolus, that CTV was cropped 3 mm from the external contour for plan evaluation and dose accumulation analysis with "virtual bolus" being applied during plan optimization to moderate the fluence in the skin region [212]. In those cases, the treating radiation oncologist omitted the use of bolus to reduce the risk of severe skin toxicity. PTVs that were located within 3 mm of the external contour were cropped back 3 mm from the skin for plan evaluation after optimization.

Daily (n = 30) cone beam computed tomography (CBCT) scans were acquired to verify patients' treatment position. The CBCT scans with a voxel size of 0.5 x 0.5 x 2 mm^3 were acquired prior to treatment using a Varian Truebeam (v2.0 or v2.5; Varian Medical Systems, Palo Alto CA, USA). For a minority of treatment fractions, CBCT scans were acquired on a Varian Clinac (v2.1) with a voxel size of 0.7 x 0.7 x 2.5 mm^3. All patients consented to their data being used for retrospective audits conform the guidelines of the local ethics committee.

4.2.2 Dose accumulation and uncertainty estimation

The total delivered dose at each successive fraction was calculated according to the DIR facilitated dose accumulation workflow previously described [213]. In summary, first a 6 degree of freedom rigid bony anatomy registration of the CBCT and the *pCT* was performed. Second, the *pCT* was deformed to match the anatomy of each daily CBCT using Varian's demons DIR implementation in SmartAdapt (SA) (v.13.6, Varian Medical Systems, Palo Alto CA, USA), which produced a deformed *pCT* (*dCT*) and a 'forward' deformation vector field (DVF). Third, the original treatment plan was recalculated on the *dCT* using the beam arrangement, monitor units and fluence maps from the original treatment plan. The resulting dose distribution was mapped back to the *pCT* space according to the true inverse DVF, producing the fraction-specific reconstructed dose. Fourth, the reconstructed dose distributions of successive fractions were accumulated. The true inverse DVF calculation and dose accumulation of the individual fraction reconstructed doses was carried out using 3D Slicer (v4.8) which is available as freeware [200,201].

The uncertainty in the dose accumulation procedure was assessed in a previous study [213] by comparing DIR facilitated dose accumulation using SA with the results of an *in silico* model based on clinically observed deformations as ground truth. These differences were separately calculated for inverse consistent and inverse *in*consistent voxels. The distinction between those voxels was made by successive application of the forward DVF and inverse DVF and classifying the net shifts larger than one dose calculation voxel as inverse *in*consistent. These results from the previous study were subsequently used in the current study to calculate the dose reconstruction uncertainties of inverse consistent u_c and inverse *in*consistent u_i voxels. The 95% level of confidence of the accumulated dose for a single voxel within structure *S* at fraction *f*, uA_f^S is:

$$uA_f^{S+}\,[\%] = \frac{\sqrt{\sum_{i=1}^{f}\left(D_i^r \,.\, u\right)^2}}{D_f^a} \times 100\% \qquad u = \begin{cases} u_c^{S+} & \text{if inverse consistent} \\ u_i^{S+} & \text{if inverse } \textit{in}\text{consistent} \end{cases}$$

$$uA_f^{S-}\,[\%] = \frac{\sqrt{\sum_{i=1}^{f}\left(D_i^r \,.\, u\right)^2}}{D_f^a} \times 100\% \qquad u = \begin{cases} u_c^{S-} & \text{if inverse consistent} \\ u_i^{S-} & \text{if inverse } \textit{in}\text{consistent} \end{cases}$$

where D_i^r are the reconstructed doses of each fraction and D_f^a is the accumulated dose for fraction *f*. The 95% level of confidence for each dose bin in the (cumulative) DVH of the accumulated dose was calculated by averaging uA_f of all voxels with D_f^a equal to or larger than that dose level.

4.2.3 Dose analysis

The accumulated dose at the end of treatment, D^{acc} was analyzed for target volumes and OARs as detailed in Supplementary Data **A.14:** Structures available for analysis for both 5 and 3 mm margin expansion plans**.** The difference between the planned dose D^{plan} and D^{acc} was defined as the change in dose, ΔD. From this point onward, DVH dose metrics of D^{plan}, D^{acc} and ΔD will be expressed using subscripts. For example, $D_{99\%}^{acc}$ refers to the minimum accumulated dose delivered to 99% of the volume. Target coverage of the PTV was assessed using $D_{98\%}$ to be consistent with both our department protocol and existing literature [23,211]. The CTV coverage was quantified using $D_{99\%}$ to enable direct comparison of our results with those reported by van Kranen et al. [142]. However, DVH metrics inherently lack spatial information regarding the dose distribution within a volume of interest [190] and

specifically for large target volumes, a clinically relevant loss in sub-volume coverage may not be detected. Therefore, CTV coverage was also assessed by recording the number of voxels where D^{plan} and D^{acc} were less than 95% of D^{presc} at the 95% confidence level. In addition, a more in-depth investigation was conducted to quantitatively assess coverage near the skin considering that target coverage was often already tight during treatment planning in this region. Specifically, the dependence of target coverage both during planning and treatment on the minimum distance between the CTV and skin was investigated. For that purpose, *local* volumes of approximately 0.5 cm^3 were defined at the point of minimum distance between the high-dose (HD) CTV and skin. A detailed methodology how the *local* volumes were constructed is provided in Supplementary Data **A.15:** Generation of local volumes.

4.2.4 Statistical analyses

The normality of the distributions of the DVH parameters D^{plan}, D^{acc} and the difference ΔD between these metrics for the two margin expansions' structures were tested using Q-Q plots and Shapiro-Wilk tests as detailed in Supplementary Data **A.16:** Normality tests. Unless stated otherwise, two-tailed paired Student's t-tests at a 5% level of significance were conducted for normally distributed metrics. The number of cases where ΔD was outside the 95% level of confidence, i.e., when the 95% CI of D^{acc} did not include D^{plan}, was defined as ψ. The number of cases where a CTV $D_{99\%}^{acc}$ was less than 95% of D^{presc} at the 95% level of confidence was defined as δ.

4.3 Results

4.3.1 Target volumes

No significant difference was observed between the coverage of target volumes for the two margin expansions at planning (Table 6) due to the minimum requirement for $D_{98\%}^{plan}$ of the PTVs in the applied planning protocol (Supplementary Data **A.9:** Planning protocol). There was also no significant difference between the observed D^{acc} during treatment for the two margin expansions, except for the low-dose (LD) target volumes LD-CTV and LD-PTV (Table 6). The average loss in target volume coverage during treatment relative to D^{plan}, ΔD, over all patients was significantly different between 5 and 3 mm margin expansions for the HD-CTV (p = 0.04), LD-CTV (p = 0.01) and intermediate-dose (ID) PTV (p = 0.03), although these differences were small. Example cases where the HD-CTV coverage during treatment was either well preserved or not are presented in Figure 4.1a and Figure 4.1b, respectively. A loss in coverage at the 95% level of confidence (i.e., *ψ)* was observed for at least 75% of all CTVs. However for only three cases, $D_{99\%}^{acc}$ was less than 95% of the prescribed dose D^{presc}. For only one case (i.e., *δ)*, this dose difference was at the 95% level of confidence.

Table 7 further details the number of voxels with less than 95% of D^{presc} for the HD-CTVs and ID-CTVs in conjunction with the observed CTV $D_{99\%}$ for individual patients. It shows that for a majority of patients, either one or both margin expansion plans included voxels that received less than 95% D^{presc} during treatment, and that this already occurred during treatment planning for many cases. In four patients, the voxels in the HD-CTV that received less than 95% D^{presc} at the 95% confidence level were located in the subdermal region (Patients 1, 10-12). For these cases, both the PTV and the CTV were not fully covered by the 95% isodose at planning except for one case (patient one; 3 mm plan) where the CTV was covered at planning. For patient three, progressive anatomical changes during treatment resulted in a small number of voxels (n = 4) at the superior side of an involved node (level 2a) receiving less than 95% D^{presc}. For patient seven, voxels in the HD-CTV receiving less than 95% D^{presc} were located in the posterior soft palate where target volume coverage was already tight during planning (95% D^{presc} isodose situated midway between the PTV and CTV). For the ID-CTV, voxels receiving less than 95% D^{presc} for patients nine, ten and twelve were located in the subdermal region.

For patient six, this underdosage was present at the left posterior aspect of the mandible where target volume coverage was tight during planning as well.

Further investigation of the relation between target coverage at the end of treatment and coverage at treatment planning showed that there was no obvious correlation between the HD-CTV $D_{99\%}^{acc}$ and the HD-PTV $D_{98\%}^{plan}$ (Figure 4.2a). There appeared to be a correlation between the HD-CTV $D_{99\%}^{acc}$ and HD-CTV $D_{99\%}^{plan}$ which were approximately equal for cases where $D_{99\%}^{plan}$ was larger than 98% D^{presc} (Figure 4.2b). For the two patients where $D_{99\%}^{plan}$ was smaller than 98% D^{presc}, both margin expansion plans exhibited a larger drop in coverage during treatment in the superficial region. There was no obvious correlation between the planned and received dose for individual voxels (Figure 4.3). This was further illustrated by the analysis of the local target coverage near the skin. Figure 4.4 shows that overall, a lower D_{min}^{plan} was obtained for the *local* CTVs that were closer to the skin. In addition, less superficial CTVs generally showed a slight increase in target coverage whereas the *local* CTVs more proximal to the skin displayed either preservation or a considerable drop in *local* coverage during treatment. These trends were very similar for both margin expansions and were independent of the magnitude of the contour change.

*Table 6: Structure-specific average and range for the planned, accumulated and Δ doses. HD = high-dose. ID = intermediate-dose. LD = low-dose. CTV = clinical target volume. PTV = planning target volume. SC = spinal cord. BS = brainstem. PRV = planning risk volume. PG = parotid gland. SMG = submandibular gland. Ips. = ipsilateral. Contra. = contralateral. D^{plan} = planned dose. D^{acc} = accumulated dose. ΔD = change in dose = $D^{acc} - D^{plan}$. $D_{99\%}$ = dose to 99% of volume. $D_{98\%}$ = dose to 98% of volume. $D_{0.1cc}$= minimum dose to 0.1 cm^3 of the volume receiving the highest dose. D_{mean} = mean dose. ψ = number of cases where ΔD was outside the 95% level of confidence. δ = number of cases where $D^{acc}_{99\%}$ was less than 95% of the prescribed dose at the 95% level of confidence.*Indicates a statistically significant difference.*

				D^{plan} [Gy]			D^{acc} [Gy]			ΔD [Gy]				
Metric	**Structure**	**n**	**PTV/PRV expansion**	**Median**	**Range**	**p-value**	**Median**	**Range**	**p-value**	**Median**	**Range**	**p-value**	ψ	δ
$D_{99\%}$	**HD-CTV**	**11**	**3 mm**	65.1	63.6 – 65.7	0.66	65.0	62.7 – 65.9	0.24	-0.3	-1.7 – 0.6	**0.04***	**9**	1
			5 mm	65.0	63.7 – 65.9		65.4	62.0 – 66.3		0.3	-2.2 – 0.7		**10**	0
	ID-CTV	**4**	**3 mm**	60.1	59.0 – 60.4	0.91	60.2	58.9 – 60.4	0.37	0.1	-0.3 – 0.2	0.44	**3**	0
			5 mm	60.0	59.6 – 60.2		60.2	59.8 – 60.6		0.4	-0.4 – 0.6		**4**	0
	LD-CTV	**12**	**3 mm**	53.5	52.3 – 54.1	0.08	53.3	52.0 – 54.4	**< 0.01***	-0.1	-1.0 – 0.3	**0.01***	**9**	0
			5 mm	53.7	52.5 – 54.6		54.1	52.4 – 55.1		0.2	-0.8 – 0.6		**10**	0
$D_{98\%}$	**HD-PTV**	**11**	**3 mm**	63.5	62.7 – 64.2	0.23	62.0	59.8 – 63.3	0.66	-1.4	-3.1 – -0.4	0.08	**11**	
			5 mm	63.4	62.8 – 63.6		62.5	59.8 – 63.2		-0.9	-3.0 – -0.3		**11**	
	ID-PTV	**4**	**3 mm**	57.6	57.1 – 58.2	0.53	56.3	55.5 – 56.8	0.12	-1.4	-1.8 – -0.9	**0.03***	**4**	
			5 mm	57.6	57.2 – 57.7		56.7	56.3 – 56.9		-0.8	-1.2 – -0.7		**4**	
	LD-PTV	**12**	**3 mm**	51.7	51.3 – 53.4	0.11	50.2	49.0 – 52.0	**0.05***	-1.5	-2.9 – -0.5	0.14	**12**	
			5 mm	52.0	51.4 – 53.5		50.6	50.0 – 52.0		-1.3	-2.2 – -0.7		**12**	
$D_{0.1cc}$	**SC**	**12**	**3 mm**	34.1	19.4 – 43.4	0.63	34.5	19.9 – 42.5	0.54	-0.1	-1.0 – 1.0	0.33	**6**	
			5 mm	36.0	23.5 – 42.3		36.1	24.1 – 41.8		0.1	-0.8 – 1.2		**3**	
	BS	**12**	**3 mm**	37.1	30.0 – 51.9	0.63	37.7	30.1 – 51.5	0.58	0.2	-2.1 – 1.7	0.41	**5**	
			5 mm	35.6	27.1 – 52.4		35.4	26.5 – 52.3		0.1	-1.4 – 1.1		**2**	
$D_{0.1cc}$	**SC PRV**	**12**	**3 mm**	38.7	21.1 – 45.2	**0.03***	38.7	21.4 – 43.9	**0.04***	-0.7	-1.6 – 1.0	0.83	**10**	
			5 mm	42.1	26.6 – 47.4		41.3	26.7 – 47.9		-0.3	-2.7 – 0.5		**7**	
	BS PRV	**12**	**3 mm**	41.5	33.7 – 52.9	**0.02***	41.3	32.6 – 52.7	**< 0.01***	0.3	-1.1 – 2.3	0.60	**2**	
			5 mm	45.3	36.5 – 60.7		45.5	35.8 – 61.1		-0.4	-0.9 – 2.6		**1**	
D_{mean}	**Ips. PG**	**14**	**3 mm**	28.0	14.1 – 43.6	**< 0.01***	28.8	14.1 – 43.0	**< 0.01***	0.4	-0.7 – 2.1	0.57	**5**	
			5 mm	31.9	16.8 – 50.0		33.4	16.7 – 49.4		0.4	-0.6 – 2.1		**5**	
	Contra. PG	**10**	**3 mm**	17.1	4.5 – 32.8	**< 0.01***	17.9	4.6 – 32.2	**< 0.01***	0.1	-0.6 – 1.2	0.71	**7**	
			5 mm	21.8	5.2 – 37.2		22.7	5.3 – 36.7		0.2	-0.9 – 1.3		**6**	
D_{mean}	**Ips. SMG**	**13**	**3 mm**	63.5	1.1 – 65.8	**0.01**	62.8	1.2 – 66.0	**< 0.01***	0.0	-1.7 – 0.7	0.17	**6**	
			5 mm	64.1	1.5 – 66.1		64.0	1.6 – 66.2		0.1	-1.8 – 0.7		**7**	
	Contra. SMG	**9**	**3 mm**	50.7	0.7 – 55.5	**0.04***	50.5	0.7 – 55.9	**0.03***	0.0	-1.1 – 1.3	0.74	**6**	
			5 mm	51.3	0.9 – 56.0		50.8	0.9 – 56.2		0.0	-0.7 – 0.8		**6**	

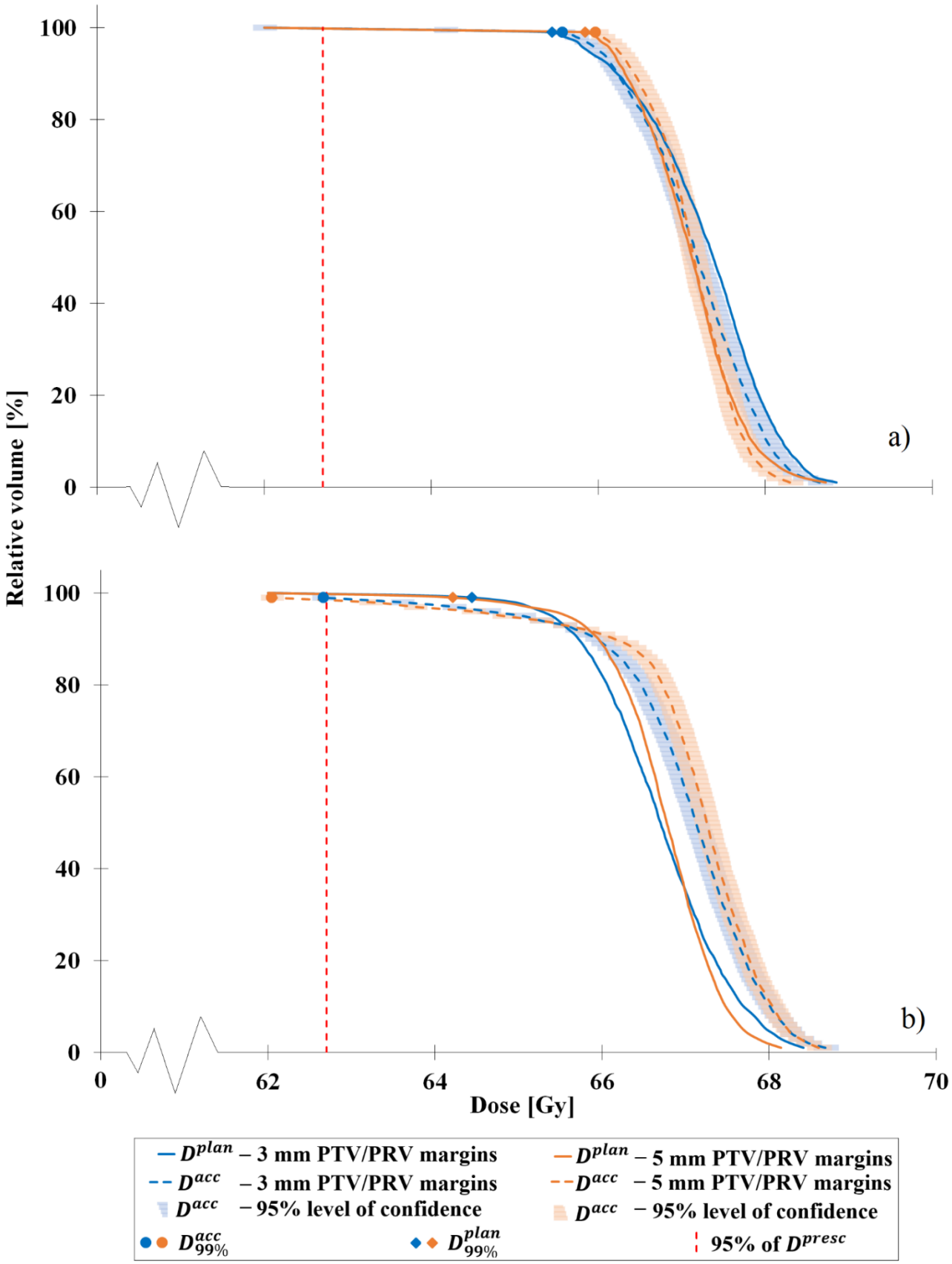


Figure 4.1: Dose-volume histograms (DVHs) of D^{plan} and D^{acc} for the high-dose clinical target volume (CTV) of patient 8 (a) and patient 10 (b). In contrast to (a), a CTV underdosage is observed for case (b) when utilizing 3 and 5 mm planning target volume (PTV) margin plans at the 95% level of confidence (b). D^{plan} = planned dose. D^{acc} = accumulated dose. $D_{99\%}$ = dose to 99% of volume. D^{presc} = prescribed dose.

Table 7: Target coverage for individual patients' high-dose (HD) and intermediate-dose (ID) clinical target volume (CTV). D^{plan} = planned dose. D^{acc} = accumulated dose. D^{presc} = prescribed dose. $D_{99\%}$ = minimum dose received by 99%

CTV	Patient	Volume [cm³]†	Margin expansion	N_{voxels}< 95% D^{presc}		$D_{99\%}$ [Gy] (95% level of confidence)	
				D^{plan}	D^{acc}	D^{plan}	D^{acc}
HD-CTV 95% D^{presc} = 62.7 Gy	**1**	**88.0**	**3 mm**		6*	64.8	**64.5** (-0.13; +0.15)
			5 mm	15*	17*	64.3	**64.1** (-0.13; +0.15)
	3	**57.1**	**3 mm**		4*	64.7	**65.3** (-0.13; +0.15)
			5 mm			65.0	**65.7** (-0.13; +0.15)
	4	**43.6**	**3 mm**			65.2	**64.8** (-0.13; +0.15)
			5 mm			65.2	**65.6** (-0.13; +0.15)
	5	**56.1**	**3 mm**			65.1	**65.5** (-0.13; +0.15)
			5 mm			65.1	**65.7** (-0.13; +0.15)
	6	**42.4**	**3 mm**			65.7	**65.3** (-0.13; +0.15)
			5 mm			65.1	**65.4** (-0.13; +0.15)
	7	**112.4**	**3 mm**		58*	64.6	**63.9** (-0.13; +0.15)
			5 mm		36*	64.6	**64.9** (-0.13; +0.15)
	8	**46.4**	**3 mm**			65.4	**65.6** (-0.13; +0.15)
			5 mm			65.8	**66.0** (-0.13; +0.15)
	9	**88.5**	**3 mm**			65.5	**65.9** (-0.13; +0.16)
			5 mm			64.6	**65.1** (-0.13; +0.15)
	10	**57.2**	**3 mm**	19*	127*	64.4	**62.7** (-0.13; +0.15)±
			5 mm	35*	175*	64.2	**62.0** (-0.13; +0.15)*
	11	**219.3**	**3 mm**	71*	77*	65.2	**65.0** (-0.13; +0.15)
			5 mm	12*	27*	65.9	**66.3** (-0.14; +0.16)
	12	**66.1**	**3 mm**	58*	99*	63.6	**62.8** (-0.13; +0.15)±
			5 mm	37*	86*	63.7	**63.0** (-0.13; +0.15)
ID-CTV 95% D^{presc} = 57.0 Gy	**6**	**78.3**	**3 mm**		4*	60.0	**60.3** (-0.16; +0.11)
			5 mm			59.6	**60.0** (-0.16; +0.11)
	9	**211.3**	**3 mm**	4*	9*	60.2	**60.4** (-0.16; +0.11)
			5 mm			60.0	**60.4** (-0.16; +0.11)
	10	**171.1**	**3 mm**		13*	60.4	**60.1** (-0.16; +0.11)
			5 mm		2*	60.0	**60.6** (-0.16; +0.11)
	12	**113.8**	**3 mm**	4*	13*	59.0	**58.9** (-0.16; +0.11)
			5 mm		2*	60.2	**59.8** (-0.16; +0.11)

† 1 cm³ includes 189 voxels * Dose less than 95% D^{presc} at the 95% level of confidence

± Dose less than 95% D^{presc} but not at the 95% level of confidence

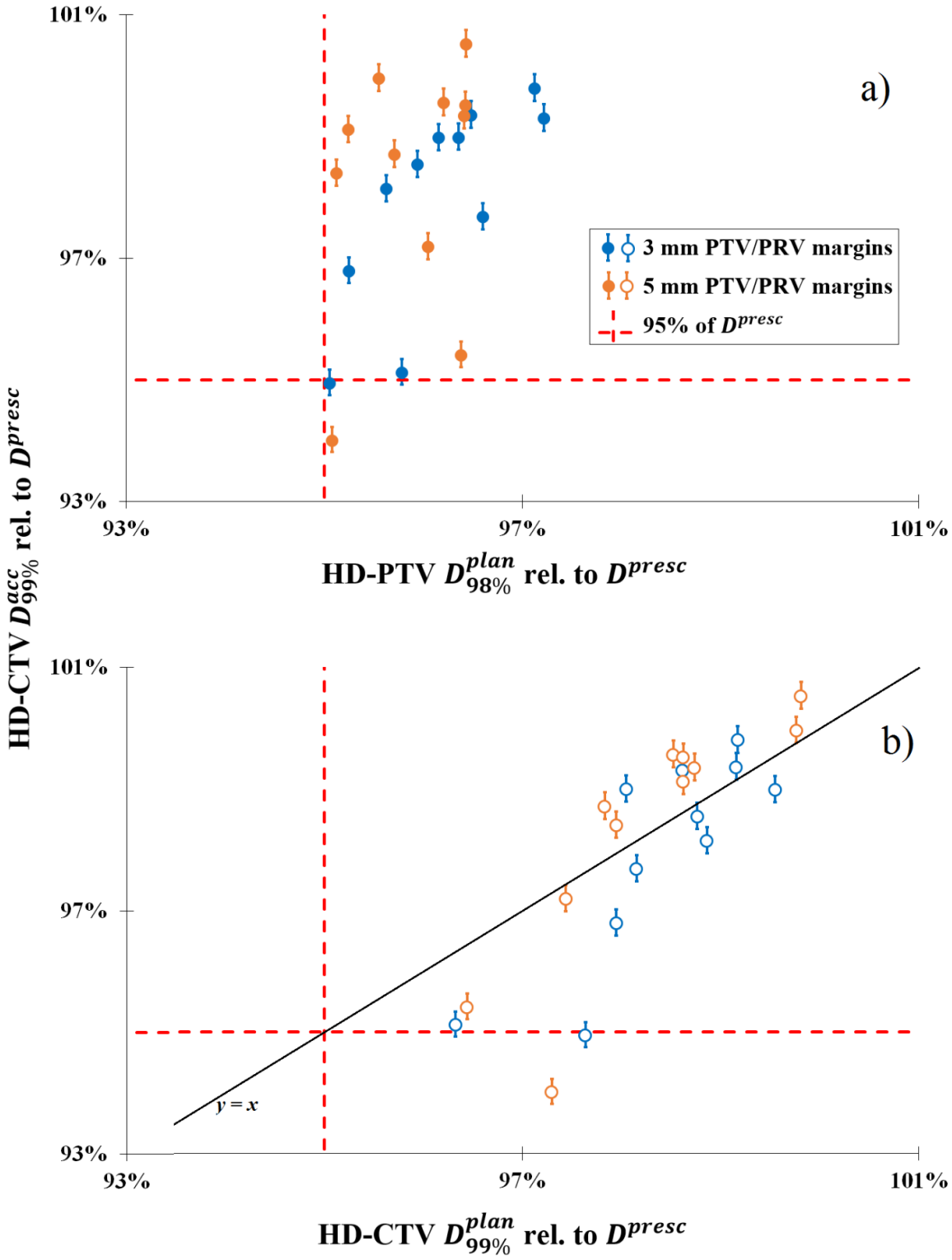


Figure 4.2: Plots of dose-volume histograms (DVH) metrics of interest for the high-dose (HD) clinical target volume (CTV) and planning target volume (PTV) when utilizing 3 or 5 mm planning target volume (PTV) and planning risk volume (PRV) margin plans. D^{plan} = planned dose. D^{acc} = accumulated dose. $D_{99\%}$ = dose to 99% of volume. $D_{98\%}$ = dose to 98% of volume. D^{presc} = prescribed dose.

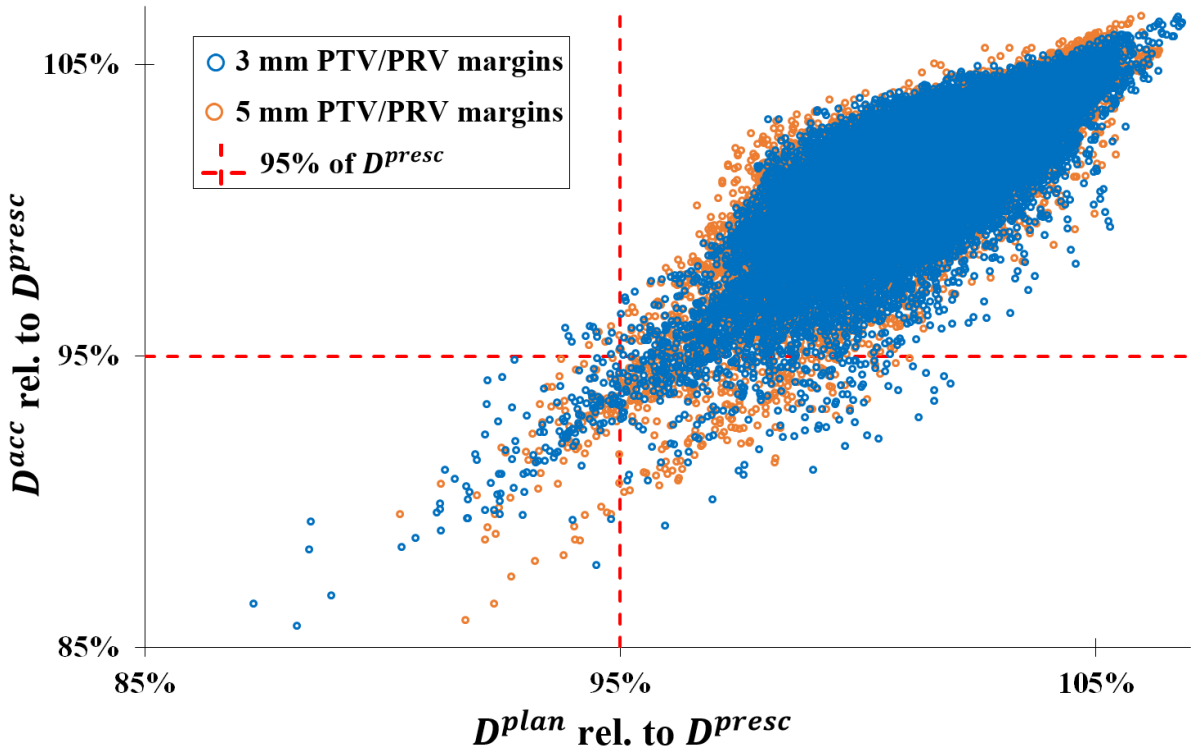


Figure 4.3: Plot of individual voxels within the high-dose clinical target volumes (CTVs) of all patients when utilizing 3 or 5 mm planning target volume (PTV) and planning risk volume (PRV) margin plans. D^{plan} = planned dose. D^{acc} = accumulated dose. D^{presc} = prescribed dose.

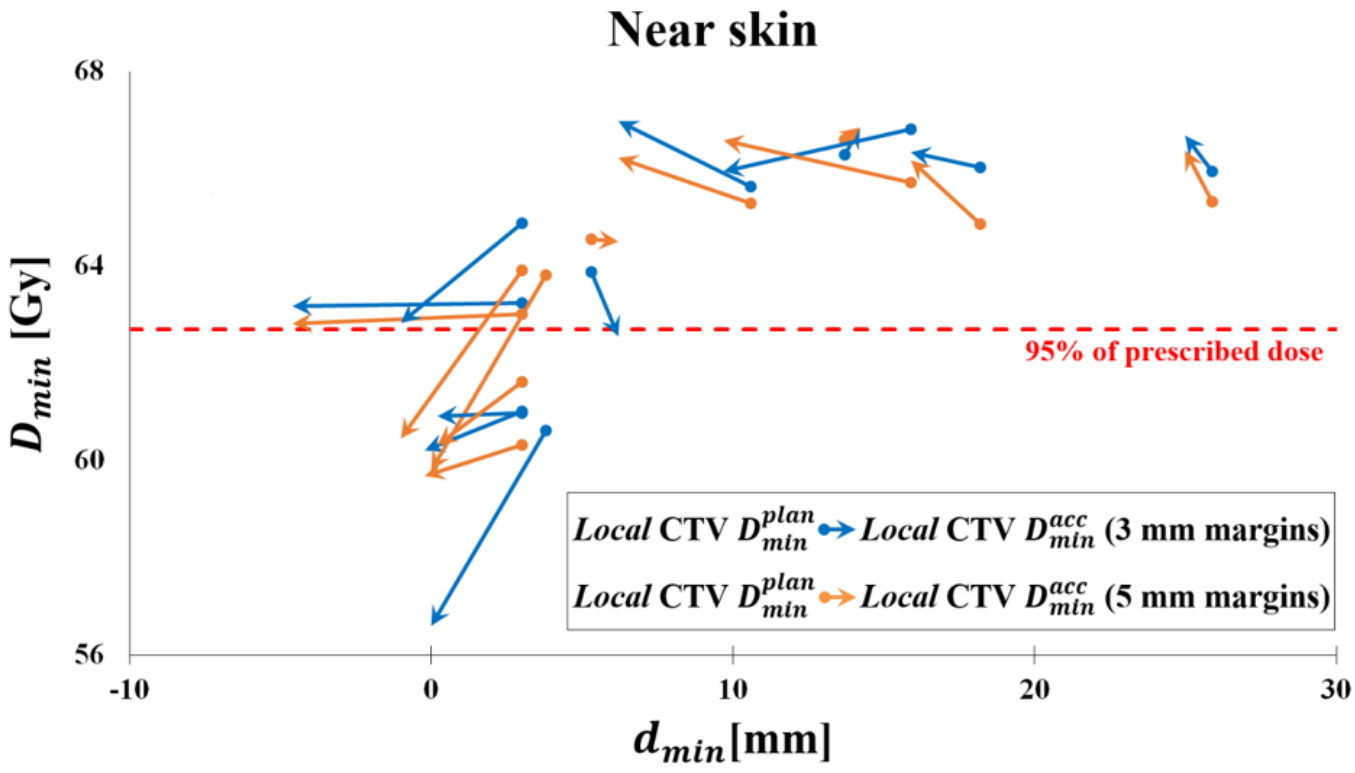


Figure 4.4: Progression of the minimum planned dose D_{min}^{plan} to the minimum accumulated dose D_{min}^{acc} for local clinical target volumes (CTVs) near the skin for 3 and 5 mm planning target volume (PTV)/planning risk volume (PRV) margin plans. D_{min}^{plan} is plotted as a function of the minimum distance between the CTV and skin. D_{min}^{acc} is plotted as a function of the minimum distance between the CTV and skin, and the local shift of the skin at the end of treatment to indicate observed anatomical changes.

4.3.2 OARs

Application of reduced PTV/PRV margins had a beneficial impact on the dosimetry for OARs (Table 6). A significantly different $D_{0.1cc}^{plan}$ was achieved for the spinal cord (SC) PRV and brainstem (BS) PRV with 3 mm margin expansions. In addition, a compromise between target coverage and maximum dose to the BS PRV due to their proximity could be avoided for one patient by using 3 mm margin expansions, whereas a higher dose than the tolerance for the BS PRV had to be accepted to obtain sufficient target coverage both for the clinical plan and for the 5 mm plan in this study. More detail on the change in dosimetry of OARs during treatment for individual cases is provided in Supplementary Data **A.17:** Individual patient OAR dosimetry. A significantly different D_{mean}^{plan} could be obtained using 3 mm margin expansions for all salivary glands. The commonly used cut-off point in normal tissue complication probability (NTCP) models [25,59] stating that the PG D_{mean}^{plan} should be preferably less than 26 Gy was achieved in 8 and 14 cases (Supplementary Data **A.1:** Patient group) for 5 and 3 mm margin expansions, respectively. The change in PG mean dose during treatment was generally less than 1 Gy and not significantly different for the two margin expansions (Supplementary Data **A.17:** Individual patient OAR dosimetry).

4.4 Discussion

The current study used DIR-facilitated dose accumulation to assess the actually delivered dose in the presence of anatomical changes, and the impact of a reduced PTV margin on the robustness of target coverage for anatomical changes. The CTV coverage $D_{99\%}^{acc}$ was found to be similar for both margin expansions and for only one case, this was less than 95% of the prescribed dose D^{presc} at the 95% level of confidence. However for many cases, the accumulated dose included a substantial volume of voxels receiving less than 95% D^{presc} independent of margin expansion. This predominantly occurred in the subdermal region.

Whether or not a PTV margin reduction is safe depends on many department-specific factors and can ultimately only be established by reviewing clinical results. Chen et al. [1,2] reported a 78% and 80% three-year LR control rate for the patient groups treated with 5 and 3 mm PTV margins, respectively (p = 0.75). Similarly, Navran et al. [3] reported that the two-year LR control rates equal to 79.2% and 79.9% for patients treated with 5 and 3 mm PTV margins, respectively, were not significantly different (p = 1.0). In contrast to the previous two groups, Franzese et al. [4] found a significant difference (p = 0.045) in LR control rates which were 87.8% and 72.6% for patients treated with 5 and 3 mm PTV margins at two-year follow-up, respectively. Unfortunately, the above clinical results studies did not have a common approach to GTV to CTV margins, treatment adaptation, and treatment plan acceptance criteria. It is therefore not possible to conclude what caused the mixed clinical results and whether a PTV margin reduction is generally safe remains equivocal. The retrospective DIR-based dose accumulation studies by van Kranen et al. [142] and Wu et al [110] investigated the impact of anatomical changes to the dose delivery during modulated HNRT for different PTV margin expansions. Both studies used DVH metrics to assess the planned and delivered dose and, well aligned with the results of the current study, both studies reported a reduction in OAR dose with reduced margin expansions and a limited change in target coverage as indicated by DVH metrics. The study by van Kranen et al. [142] which used dose accumulation based on DIR of daily CBCT images reported a CTV $D_{99\%}^{acc}$ less than 95% of D^{presc} for only 1/19 and 2/19 cases for 5 and 3 mm PTV margin expansions,

respectively. The study by Wu et al. [110] used dose accumulation based on weekly re-CTs included eleven patients. This study concluded that the coverage of the CTVs at the end of treatment was not affected by anatomical changes based on the ratio of multiple DVH metrics relative to the corresponding planning metrics. It should be noted that the current study and those by van Kranen et al. [142] and Wu et al. [110] may be restricted by limitations of DIR in the presence of mass and density changes [139,143,166–169]. The impact of this limitation is partially included in the dose accumulation uncertainty estimate that was used in the current study [213]. In addition, the current study and similar retrospective studies [110,142] only evaluated the plan robustness for anatomical changes during treatment. A full PTV margin estimate should include all geometrical uncertainties as sufficiently described in literature [35,36,170]. Due to the absence of a golden standard for delineation accuracy and limited knowledge on microscopic tumor spread for different tumor types [37,214,215], the only way to fully assess the validity of applied treatment margins is by pattern of failure studies or a review of clinical results as exemplified in refs [1–4] which must include consistent volume definitions [34,216]. However, we recommend that more comprehensive information is provided in review of clinical results studies including a voxel-based analysis of the actually delivered dose.

The initial analysis of target coverage based on DVH metrics in the current study suggested that 5 and 3 mm PTV margin plans have similar target coverage with only one 5 mm treatment plan where the target coverage was below the rejection criterion $D_{99\%}^{acc} < 95\%\ D^{presc}$ at the 95% level of confidence (Table 6). However, a full analysis of the dose delivered to each voxel (Table 7) highlighted potentially clinically significant underdosage in many cases both for 3 mm and 5 mm margin expansions. DVH metrics inherently lack spatial information regarding the dose distribution within a volume of interest [190] and commonly refer to percentages instead of absolute volumes. Specifically for larger target volumes, underdosage of the high- and intermediate-risk target volumes may not be detected by using DVH metrics as criterion. Therefore, a voxel-specific analysis of the entire target volume as is presented in the current study is preferred.

For the majority of cases with voxels receiving less than 95% D^{presc}, these voxels were located in the dermal region. Pattern of failure studies in literature that did explicitly document the locations of marginal loco-regional recurrence (LRR) indicated that these occur relatively frequently in the subcutaneous/dermal region. Specifically, subcutaneous/dermal recurrences constituted 45% (4/9) [1], 60% (3/5) [217], 100% (1/1) [218] and 22% (2/9) [219] of the marginal LRRs in these reports. No dermal recurrences were observed for the patients included in the current study (median follow up 32 months, range 7 – 44 months). Further retrospective review of the study cases by the most senior clinician of our department involved in HNRT revealed that in some cases the observed underdosage of specific voxels in this study could potentially be clinically relevant. However, in most cases the CTV contour near the skin might have been slightly too generous with contour extension beyond the platysma and the observed underdosage of specific voxels in this study might not be clinically relevant for those cases. It should be noted that a slightly less generous CTV contour would also have resulted in a different treatment plan with likely one or more locations presenting a compromise in target coverage. As per the current departmental procedure, this study applied "virtual bolus" during plan optimization when CTVs extended into the dermal region [212] to moderate the fluence in the skin region and reduce the risk of severe skin toxicity. This approach commonly results in tight target coverage near the skin but is generally accepted by the treating clinician as compromise rather than using bolus on treatment unless the skin itself is at risk.

For 3D-conformal radiation treatments, the accuracy of the accumulated skin dose is limited by the uncertainty in the calculated dose in the buildup region [220]. However for VMAT treatments in this study, the dose to skin could be regarded as exit dose because the majority of dose to the skin was delivered through medial beam angles. Using the methodology previously developed for time-resolved point dose QC to assess the TPS calculated dose per control point [221], it could be demonstrated that ~80% of the dose delivered to 40 superficial HD-CTV points with $D^{acc} < 95\% D^{presc}$ was deposited at beam angles where the effective depth was larger than or equal to 4 mm. More detail on this analysis is provided in Supplementary Data **A.18:** Time-resolved point dose QC analysis. In general, full scatter conditions can simply be achieved by immobilization devices such as radiation masks 'behind' the

patient. Therefore, the exit dose follows the normal percentage depth dose curve and can be accurately calculated [220,222,223]. The accuracy of the skin dose calculation could also be limited by uncertainty in the Hounsfield units (HUs) of the most superficial voxels within the body contour [224]. Supplementary Material **A.19:** HU override analysis details the results of planning simulations using an HU override equal to the average HU value ± 100 for these voxels. The HD-CTV doses were assessed for both HU values and the largest resulting 95% percentile range of the difference was ~1% D^{presc}. It is therefore reasonable to assume that any uncertainty in the HU of superficial regions has minimal impact on the observed values of HD-CTV D^{acc}, which were as low as ~85% D^{presc} (Figure 4.3).

In the current study, loss in CTV coverage was very patient specific and appeared to be independent of margin expansion. Specifically for superficial HD-CTVs, the loss in *local* target coverage occurred more frequently in regions near the skin, but appeared not to be correlated with the magnitude of anatomical changes. In general, multiple mechanisms play a role in target coverage differences between planning and treatment. The interplay between local fluence, dose delivered per gantry angle, contribution of scatter and anatomical changes prohibited a clear identification of risk factors for individual patients that may be prone to a loss in target coverage during treatment. Nevertheless, it was clear that target coverage at treatment planning was an important factor considering that changes in patient anatomy on treatment are more likely to result in a decrease in target coverage when it was already marginal or compromised during the treatment preparation stage. It is recommended to assess target coverage using a voxel-based approach and not exclusively rely on DVH metrics. In principle this is not different from a slice-by-slice review of the 3D dose distribution before acceptance at treatment preparation by the treating radiation oncologist. However, manual review of the delivered dose after each fraction would be very laborious and automated processing of results using a voxel based analysis would be preferable for efficiency reasons.

In addition to increased sparing of salivary glands as described by many other studies [58,110,142], this study also highlighted the increased avoidance of the critical OARs with reduced PTV margins. As the

minimum distance between the HD-CTV and BS decreased, a higher D_{min}^{plan} to the *local* PTV was generally achievable with 3 than with 5 mm margins (Supplementary Data **A.17:** Individual patient OAR dosimetry). Comparably, as the minimum distance between the HD-CTV and BS decreased a lower D_{max}^{plan} to the *local* PRV was generally achievable with 3 mm margins (Supplementary Data **A.17:** Individual patient OAR dosimetry). The compromise to the local BS PRV D_{max}^{plan} as required for the 5 mm margin plan for patient ten was not necessary for the 3 mm plan.

In summary, a PTV/PRV margin reduction from 5 to 3 mm itself did not change the plan robustness for anatomical changes during treatment. However, a considerable loss in CTV coverage was observed for some patients irrespective of the margin expansion. Although this was very patient specific, it was specifically observed when target coverage at treatment planning was already tight. Patient-specific verification of dose delivery during treatment is therefore recommended, for instance using the DIR-facilitated dose accumulation analysis presented in the current study.

5 DISCUSSION

5.1 Overview

The motivation for conducting the body of research detailed in this dissertation was to demonstrate that a safe PTV margin reduction from 5 to 3 mm is feasible by developing a quality management "safety net" for patients who may be at risk of unacceptable loss in target coverage and/or overdosage to normal tissues. The use of a reduced 3 mm PTV margin is likely to improve patient toxicity results [1–4]. An emphasis on *safe* PTV margin reduction is particularly warranted considering that an equivalent LR control rate was not maintained for one [4] of three groups that have clinically implemented reduced PTV margins [1–4], and that the frequency of loss in target coverage during treatment may slightly increase with reduced mm PTV margins [142]. It is important to optimise the PTV margin so appropriate target coverage is achieved while simultaneously ensuring the margins are not overly generous resulting in unnecessary toxicity for patients.

PTV margins do not standardly account for non-rigid anatomy changes (e.g., changes in patient pose, weight loss, tumour response, OAR shift, OAR shrinkage) commonly observed during HNRT which can undermine the planned dose objectives. Therefore, a loss in target coverage could occur more often and may become more relevant with reduced mm PTV margins. The premise of developing the quality management "safety net" in the research described in this dissertation was to enable the monitoring and quantification of a patient's dosimetry in the presence of anatomical changes. Monitoring the delivered dose during treatment would allow detection of unacceptable loss in target coverage and/or overdosage to OARs. Ultimately, identifying patients at risk would facilitate timely and appropriate ART. Broadly, the "safety net" was required to (*a*) quantify and separate the different modes of non-rigid anatomical change; and (*b*) quantify the dosimetric impact of non-rigid deformations. The studies included in the

research described in this dissertation were undertaken to build the quality management "safety net", and resulted in a series of papers for publication which were reproduced in *Chapters Two, Three* and *Four* (at the time of this dissertation's submission *Chapter Two* has been published, *Chapter Three* has been accepted and is an article in press and *Chapter Four* is in peer-review). The *Sections* 5.2-5.4 of the current *Discussion* chapter detail the "safety net" requirement that each study aimed to achieve and the rationale for conducting each study in the context of the wider body of literature. In addition, further considerations that are relevant to each study which were not provided in the corresponding discussion *Sections* (i.e., 2.4, 3.4 and 4.4) are put forward. *Section* 5.5 presents a collective summary of the individual studies to show how they constitute a coherent body of research and the current *Discussion* chapter concludes with a consideration of future perspectives.

5.2 Monitoring anatomical changes of individual patients using statistical process control during head-and-neck radiotherapy

The study described in *Chapter Two* aimed to develop a method to quantify and separate the different modes of non-rigid anatomical change that can occur during treatment, the "safety net" requirement (*a*). A large number of previous studies have successfully quantified non-rigid deformations in HNRT [78,81–84,91,95,98,225,226]. These studies include the quantification of daily positioning inaccuracies [81–85], investigations of the change in volume and neck diameter of various HN sub-regions as a surrogate of weight change [78,93,95,226], and investigations of structures' volumetric changes [79,91,98]. However, the analyses in these studies were either carried out manually which would make their clinical implementation impractical from an economic and efficiency perspective, or did not provide a method of detecting systematic changes. SPC (*Sub-section* 2.2.5) is widely used in manufacturing as a quality management tool and recently, its use in radiotherapy for quality assurance has been discussed [188]. Control charts, also known as process-behaviour charts, are frequently used to examine and monitor process time trends [189,227]. Natural variation will occur in even the most well controlled processes. Discrimination of natural and exceptional variation is governed by a single interpretation rule for an EWMA SPC chart, i.e., if a point falls outside the control limits this is

interpreted as exceptional variation and the cause should be investigated. The unambiguous instructions surrounding control chart initialisation and interpretation allow time trends to be reviewed in a standardised way.

The feasibility study in *Chapter Two* was the first study to investigate the suitability of DIR and EWMA SPC charts to assess trends for individual patients during HNRT. Specifically, DIR and EWMA SPC charts were used to objectively quantify, monitor and separate individual patients' non-rigid deformations. A standardised approach to quantify patient deformations with an estimated robustness derived from comprehensive uncertainty analyses was utilised. Trends in patient positioning reproducibility occurring before the end of treatment week four were confirmed in 24% of cases when SPC process limits were used, whereas absolute BA structure deviations could be confirmed in 82% of cases when an *a priori* 2 mm clinical limit was used. SPC process limits were useful to detect trends of changes of the treatment-region volume, PGs shrinkage, and PG shifts occurring before the end of treatment week four which could be confirmed in 90% of cases.

The main challenge encountered in the study in *Chapter Two* was defining the exact moment a patient positioning deviation trend exceeded the SPC control limits at a certain statistical confidence interval. This was particularly applicable to shallow trends with narrow process limits in spite of their sub-millimetre accuracies. For these cases, the relatively large uncertainties in both the DIR derived EWMA statistics and the EWMA process limits precluded trend confirmation at a 90% confidence level. The main contributors to the total uncertainty in the DIR derived EWMA statistics were the uncertainty in the DIR algorithm, and the uncertainty when copying structures to the CBCT from the dCT (Supplementary Data **A.2:** Uncertainty analysis). The accuracy of the DIR algorithm will likely be improved by the application of local rigidity constraints to BA structures to reduce the impact of "creep" [197,198] while an improved level of uncertainty when copying structures to the CBCT from the dCT is expected with improved image resolution of the planning CT, particularly in the z direction. The uncertainty in the EWMA process limits was caused by the uncertainty in the standard deviation of the reference period (uncertainty in σ, equation 4 page 79) only, because the other variables determining

the process limits are constants. The uncertainty in the EWMA process limits was therefore directly proportional to the variation in the reference period. Efforts to reduce the uncertainty in the EWMA process limits could focus on defining the optimal length of the reference period which is likely to be structure- and patient-specific. The study in *Chapter Two* used the first week of treatment, fractions 1-5, as the reference period length to represent a "stable" process. However, the first week of treatment may not always appropriately represent a stable process as potential trends for some structures and for some patients could already be occurring. Investigating the methods that were proposed in the current paragraph to reduce uncertainty and optimise the EWMA reference period may indeed develop structure-specific clinical limits which may then demonstrate the efficacy of SPC to timely indicate ART.

Considering the difficulty with shallow trends and narrow SPC process limits as described in the preceding paragraph, the study in *Chapter Two* found it more efficient to apply a clinical limit to 'flag' patient positioning deviations that may have a clinical impact during treatment. In that study, a generic *a priori* 2 mm clinical limit was used for all BA structures. However, on clinical implementation, the clinical limit should ideally be individually set for each BA structures' deviations and for each patient based on expected impact to target coverage and/or OAR sparing. To further clarify, it helps to realise that BA deviations are essentially surrogates for the non-rigid positioning of target volumes and the extent to which a target volume may abut or even contain a BA structure needs consideration. For example, a target volume immediately adjacent to C1-C3 that also encapsulates the entire mandible is likely to have a clinically relevant change in dosimetry as the mandible deviation reaches the size of the PTV margin expansion. Whereas for a target volume immediately adjacent to C1-C3 extending only midway to the mandible, the same mandible deviation magnitude is less likely to cause a clinically relevant change in dosimetry.

The study in *Chapter Two* successfully developed a method to quantify and separate the different modes of non-rigid anatomical change. However, non-rigid deformation trends exceeding an SPC process or clinical limit themselves were not indicative for the clinical relevance and required additional

interpretation rules. *Chapter Two* set the platform required to investigate the dosimetric impact of the observed non-rigid deformations in an effort to show the efficacy of SPC to timely indicate required ART.

5.3 Quantifying the dose accumulation uncertainty after deformable image registration in head-and-neck radiotherapy

DIR-facilitated dose reconstruction and accumulation allows estimation of the actually delivered dose in the presence of anatomical changes during HNRT [139]. DIR-facilitated dose accumulation was therefore identified as a tool that could quantify the dosimetric impact of non-rigid deformations, the "safety net" requirement (*b*). However, the validation of DIR for this purpose is challenging because a comprehensive ground truth when dealing with patient data does not yet exist [130]. In the absence of a comprehensive ground truth, a number of studies [136,145,164,228–230] have validated the accumulated dose using physical or simulated phantoms with known DVFs. However, the reference deformations in these studies were not always clinically realistic. Other groups have indirectly estimated the uncertainty in the accumulated dose by modelling DIR uncertainties as multivariate Gaussian distributions [231,232]. As discussed in the DIR for dose accumulation review of Chetty et al. [139], specific aspects of DIR-facilitated dose accumulation that can impact the accumulated dose have been investigated by various authors, i.e.,: the choice of DIR algorithm and parameters [145,162,230,233], the ROI selected to generate the DVF [228], the dose grid resolution [234], the inverse consistency of the DVF [159,165] and the principles of mass and energy conservation [229,235]. Current literature encourages research that validates DIR-facilitated dose accumulation in the setting of clinical realism [139], in particular for patients exhibiting mass/density changes during treatment [169].

As it was considered whether it was possible to utilise DIR-facilitated dose accumulation to assess the delivered dose in HNRT using the DIR application available at the WBCC, it was recognised that the factors which may impact the accumulated dose mentioned in the preceding paragraph are algorithm, treatment site and potentially department specific, and should be explored. To the best of the author's

knowledge, no previous study has evaluated the uncertainty in the proposed dose accumulation clinical workflow with subsequent incorporation in DVHs. The study in *Chapter Three* was conducted to address this knowledge gap, and the general recommendation of the literature mentioned in the previous paragraph [139,169].

The retrospective study in *Chapter Three* utilised *in silico* deformations based on clinically observed anatomical changes as ground truth to evaluate the dose accumulation uncertainty. In order to assess the structure-specific uncertainties of the demons DIR algorithm for inverse consistent and inverse *in*consistent voxels, the reconstructed dose based on demons DIR was compared to a ground truth reconstructed dose distribution based on B-spline DIR. The inverse consistency of voxels was assessed by determining their net displacement after successive application of the forward and backward DVF. Over all patient structures included in this study, 98.5% of voxels were inverse consistent with a 95% level of confidence range for dose reconstruction of a single fraction relative to their planned dose equal to [-2.3%; +2.1%], [-10.2%; +15.2%] and [-9.5%; +12.5%] for target structures, critical OARs and non-critical OARs, respectively. As per Equation 7 (page 108), the uncertainty in the accumulated dose for these structures was [-0.4%; +0.4%], [-1.9%; +2.8%] and [-1.7%; +2.3%] at the end of treatment, respectively. Inverse *in*consistent voxels generally showed a higher level of uncertainty (Table 5). Quantification of the uncertainty in DIR-facilitated dose accumulation allowed incorporation of a 95% confidence interval in DVHs. The methodology developed in *Chapter Three* can be used to prospectively assess the adequacy of target coverage during treatment and decide whether a re-plan is necessary in an objective manner. Furthermore, the methodology described in *Chapter Three* can also be used to validate reconstructed and accumulated doses acquired with other DIR implementations and imaging schemes.

Currently, the suitability of DIR for accurate and precise dose accumulation is debated in the literature [139,143,166–169,236] which may be contributing to the absence of a large scale clinical implementation. The primary limitation of most currently available DIR algorithms causing the debate is whether DIR-facilitated dose accumulation is appropriate in the presence of mass/density changes

occurring during treatment. This is particularly applicable to regressing tumours but also some normal tissues, such as the PGs where a significant reduction in volume is often observed during HNRT [79,207]. Historically, intensity-based DIR algorithms do not have provisions to account for loss of tissue [167,168] and generally attempt to align the border of the original structure to the border of the structure that has partially dissolved/shrunken during treatment by "squeezing" voxels together. Consequently, two different voxels of the original structure may both be directed to a single voxel of the dissolved/shrunken structure. In that case, the reconstructed dose in the partially dissolved/shrunken structure is imposed back on the initial structure, irrespective of how the structure has actually changed. This would impose that same dose to all voxels of the initial structure which would create energy, and violate the principle that energy should be conserved [167,168]. In addition to the principle of energy conservation limitation, DIR-facilitated dose accumulation does not standardly consider individual cells' response to radiation exposure [139]. There are two modes of cellular tumour regression that are important in this context: elastic and inelastic [237]. Elastic regression refers to the process whereby all tissue (tumour and normal) retreats concentrically with the tumour border including any subclinical microscopic malignant disease [237]. Inelastic regression is the case where the tumour boundary retreats concentrically but independently of the surrounding normal tissue, potentially leaving pockets of subclinical microscopic malignant disease [237]. An example of inelastic regression can be observed for HN SCC patients with extracapsular spread present in metastatic lymph nodes [238,239]. For both elastic and inelastic regression modes, DIR algorithms with a strict one-to-one voxel map (as a surrogate for cell-to-cell) functionality that also facilitates tissue loss will likely provide a better estimate of the accumulated dose. The appropriateness of DIR-facilitated dose accumulation in the presence of mass/density changes specifically on the study in *Chapter Three* was discussed in *Section* 3.4 where it was concluded that the impact of volume changes on the dose reconstruction accuracy was at least partially included in the uncertainty estimate. This conclusion was appropriate considering that *in silico* deformations based on clinically observed anatomical changes were used as ground truth. The workflow for dose accumulation first described in *Chapter Three* was used in *Chapter Four* to objectively assess the adequacy of target coverage during treatment for a reduced PTV margin setting.

5.4 Individual patient dose accumulation assessing the validity of reduced PTV and PRV margin plans in head-and-neck radiotherapy

The study described in *Chapter Four* utilized DIR-facilitated dose accumulation to objectively assess the adequacy of target coverage, critical OAR avoidance and non-critical OAR sparing when either 5 or 3 mm PTV and PRV margins were applied and as a result, the dosimetric impact of the non-rigid deformations was investigated which was the "safety net" requirement (*b*). To be consistent with the scarce body of literature regarding the robustness of treatment plans for anatomical changes during HNRT when margins are reduced [110,142], target coverage, critical OAR avoidance and non-critical OAR sparing was first quantified with DVH metrics. However, DVH metrics inherently lack spatial information regarding the dose distribution within a volume of interest [190] and for large target volumes, a clinically relevant loss in local sub-volume coverage may not be detected. Therefore, a more in-depth voxel-specific investigation was conducted. To the best of the author's knowledge, no previous study has reported on the application of a DIR-facilitated dose accumulation voxel-specific analysis for a reduced PTV/PRV margin setting. Using DVH metrics and the 95% C.I. derived as described in Chapter Three, a statistically significant underdosage (defined by a CTV $D_{99\%}^{acc}$ being less than 95% of D^{presc} at the 95% confidence level) was observed in one case only while for two other cases, the underdosage was borderline significant. However a voxel-based analysis showed that for many cases, the accumulated dose included a substantial volume of voxels receiving an underdosage. Loss in target coverage was independent of margin expansion and very patient specific with tightness of target volume coverage at planning was a common factor leading to underdosage. While used retrospectively in the study in *Chapter Four*, the presented dose accumulation analysis is primarily intended for prospective clinical application.

As introduced in *Section* 1.7 classical PTV margin recipe calculations [35,36] yield CTV to PTV margin expansions of 6-9 mm, even when non-rigid deformations are ignored and low estimates for the uncertainty in target volume delineation are used. Despite this, PTV margins of 5 mm are commonly applied to establish target coverage [170,171] and high two- and three-year local tumour control rates

up to 78-88% are commonly reported [1–3,53,62]. Furthermore, the studies of Chen et al. [1,2] and Navran et al. [3] have clinically implemented reduced PTV margins from 5 to 3 mm and reported equivalent two- and three-year LR control rates and overall survival between the two margin expansions. It should also be mentioned that most treatment failures occur in field while recurrences within or near the PTV margin are relatively rare [172,173]. Specifically in HNRT there are a numbers of reasons that might explain the contradiction between 6-9 mm PTV margin calculations [170] and favourable outcomes using a 3-5 mm PTV margin: (*1*) The margin around the GTV which defines the CTV and accounts for microscopic spread might be too generous [214,215]. A lack of consensus on the appropriate CTV expansion exists in the literature [32–34,218,240,241]; (*2*) PTV margin recipes do not account for the presence of an intermediate-risk PTV in HNRT which usually embeds the high-dose PTV and primary tumour. This intermediate-risk PTV is irradiated to a lower dose level and provides a dose bath which may partially compensate geometrical misses; (*3*) As noted by Van Kranen et al. [142], a patient's deformations could shift a CTV into an adjacent PTV in cases of multiple CTVs or concave target volumes; (*4*) Volume loss of the treatment region (e.g. weight loss, tumour and/or salivary gland shrinkage) could move CTV boundaries further into the high-dose region; (*5*) The planned dose distribution does not conform perfectly to the PTV, i.e., the volume treated to 95% of the prescribed dose is larger than the PTV contour. The latter can be observed in Figure 1.4 where the 95% isodose colourwashes do not conform perfectly to the high-dose PTV contours shown in red.

Treatment plan optimisation and evaluation is typically directed by constraints and objectives in order to achieve satisfactory doses to targets and normal tissues [45]. The constraints and objectives are generally based on models of TCP and NTCP. These models are themselves normally based on treatment outcomes with dose and volume data from the *planned* dose distribution [114]. However as well-documented in the literature [58,80], the planned dose is not always equal to the dose that is actually delivered due to changes in a patient's positioning and/or anatomy during treatment. Consequently, the constraint and objective thresholds may not be appropriate benchmarks to assess the actually delivered dose or to determine those patients for treatment adaption. Accurate and precise DIR-facilitated dose accumulation could be used to quantify the actually delivered dose which may in turn

be used to update control and complication models [242]. In agreement with the literature [58,80], the results in *Chapter Four* (**Error! Reference source not found.**) found a significant difference between the planned and actually delivered doses in the majority of cases. However, the change in coverage in terms of absolute dose was relatively small. It is therefore unlikely that NTCP models can be updated in a pragmatic manner based on results such as those in *Chapter Four* considering the uncertainty intervals in NTCP models [25,59] are often larger than the change in coverage that was observed. Using the PGs as an example, the study of Hunter et al. [60] found differences between planned and delivered D_{mean} doses (-4.9 to +8.4 Gy, median difference +2.2 Gy) in the majority of PGs were small relative to the standard deviation of the dose-saliva data, suggesting that ART is not likely to gain measurable salivary output improvement in most cases. These differences were observed at first treatment, indicating potential benefit for more complex setup corrections or adaptive interventions in the minority of patients with large deviations detected early by CBCT. It is likely than a more effective gain in PG function after HNRT will be achievable with margin reduction as opposed to ART. In the results in *Chapter Four* the commonly used cut-off point in NTCP models [25,59] stating that the PG D_{mean}^{plan} should be preferably less than 26 Gy was achieved in 8 and 14 cases (Supplementary Data **A.1:** Patient group) for 5 and 3 mm margin expansions, respectively.

It should be realised that the study in *Chapter Four* as well as the two comparable studies that have been previously mentioned [110,142] evaluated treatment plan robustness in the presence of anatomical changes only, and did not consider other treatment related uncertainties, such as geometrical accuracy of the treatment machine and geometrical uncertainty of the target volume definition. In contrast to the latter two studies [110,142], the study in *Chapter Four* evaluated the delivered dose to the CTVs and PTVs. Considering that the application of a PTV margin is to ensure the CTV receives adequate dose in the presence of all treatment related uncertainties it was relevant to assess both CTV and PTV coverage to also account for the additional treatment related uncertainties.

Chapter Four was successful in quantifying the dosimetric impact of non-rigid changes in patient anatomy and did so by assessing target coverage, critical OAR avoidance and non-critical OAR sparing when either 5 or 3 mm PTV and PRV margins were used on a patient-specific basis. The results in *Chapter Four* indicated that PTV/PRV margin reduction did not significantly reduce CTV coverage during treatment. However, patient-specific verification of the dose delivered to each voxel was recommended, irrespective of margin expansion, considering that assessment of the target coverage using DVH parameters only was not sufficient.

5.5 Collective summary

The three studies in this dissertation were successful in developing the quality management "safety net" by demonstrating the feasibility of prospectively applying the tools to monitor non-rigid anatomical changes and delivered dose. More specifically, the study in *Chapter Two* presented methodology to detect systematic anatomical changes occurring during treatment and the studies in *Chapters Three* and *Four* provided a means to accumulate dose after each fraction to decide whether treatment adaption is required. It was initially intended to further extend the "safety net" by directly linking the results in *Chapters Two* and *Four* in an effort to generate generic guidelines to select and/or predict patients for timely and appropriate treatment adaption using SPC. Objective and generally applicable guidelines to select patients who will benefit from treatment adaptation are generally lacking [79,115] and while some studies have investigated the appropriate timing of a scheduled treatment adaption [127,243], commonly the decision to adapt treatment is based on manual ad hoc IGRT review in which treatment staff qualitatively address patient deformations (weight loss, tumour regression, positioning reproducibility, etc.) [93,119–121]. However considering that loss in target coverage was found to be very patient specific in the results in *Chapter Four*, the development of generic guidelines for timely and appropriate treatment adaption using SPC could not be demonstrated. The following paragraph details the ambiguity that was observed when attempting to directly link the results in *Chapters Two* and *Four*.

In total, eight patients had an underdosage (***Error! Reference source not found.***; *Section* 4.2.1) which was defined by an individual CTV voxel's accumulated dose being significantly less than 95% of D^{presc}. The underdosages for five of these patients occurred in the subdermal aspect of the CTV region while the remaining three underdosages which were non-subdermal occurred in the posterior soft palate, the superior aspect of a level 2a node and the left anterior aspect of the mandible. In an effort to develop the generic guidelines to select patients for ART, the SPC trends for these eight patients were further scrutinised. A systematic change in the neck volume occurring before the end of treatment week four was confirmed in seven patients. However, these seven patients only included two patients who had a subdermal underdosage and two patients who had non-subdermal underdosage. Therefore, approximately 50% of systematic changes in the neck region volume occurred independently of an underdosage. To further illustrate the ambiguity when attempting to establish a potential correlation between underdosage and treatment volume change, it should be noted that 40% (6/15) of the HD or ID-CTVs already had an underdosage at planning (***Error! Reference source not found.***) for at least one of the margin expansion plans. In addition, despite patient one exhibiting the largest local contour reduction of all patients, the neck region volume of patient one remained 'in-control' from an SPC perspective throughout treatment. Further investigation found wider process control limits for patient one compared to the other patients, owing to a larger variation in the reference period. Absolute deviations for those BA structures most proximal to the underdosed CTVs were also investigated. Three of the patients with subdermal underdosage had a BA deviation exceed the 2 mm clinical limit before the end of treatment week four for their most proximal BA structure (mandible or hyoid). However, the remaining two patients' deviation of their most proximal structure, C3-C5, did not exceed the clinical limit. While the results for the former three patients are encouraging it should be stated that 67% (2/3) of the patients who did not have a HD-CTV underdosage had their most proximal BA structure's deviation (mandible) exceed the clinical limit. Similarly ambiguous results were obtained for the three patients with non-subdermal underdosages where only 33% (1/3) of the most proximal BA structure's deviation exceeded the clinical limit. The feasibility of using SPC detected systematic changes in anatomy as a risk factor for CTV underdosage was unable to be demonstrated in a consistent and generally applicable manner.

The patient-specific nature of loss in target coverage hampered the development of generic guidelines for ART based on changes in patient positioning or anatomy changes. However, the DIR-facilitated dose accumulation approach that was developed did provide a means to objectively decide whether ART is necessary for individual patients. There are two modes of ART that could apply the developed dose accumulation methodology: online and offline [114,244,245]. An online ART approach could evaluate the delivered dose and correct the dosimetric impact of complex anatomical changes by re-optimising the plan on the dCT of the pre-treatment CBCT with subsequent delivery. However, the burdensome workload [246] and technological requirements [115] make it unlikely that such an online ART will be implemented on a large scale in the near future. Further discussion of online ART is provided in *Sub-section* 5.6.2 below. In contrast to online ART, it is possible that an offline ART procedure based on the DIR-facilitated dose accumulation methodology demonstrated in *Chapters Three* and *Four* could be implemented in the near future. Retrospective dose reconstruction would be conducted at the end of each treatment fraction with subsequent evaluation of the accumulated dose. Evaluation of the accumulated dose with concurrent consideration of the anticipated dose that will be delivered in the remaining treatment fractions would then facilitate the decision to adapt treatment. The plan re-optimisation could be conducted on the dCT with subsequent delivery occurring as soon as quality assurance protocols have been completed.

5.6 Future perspectives

The research described in this dissertation aimed to demonstrate that a safe PTV margin reduction from 5 to 3 mm is feasible at the WBCC. In addition, the research aimed to contribute to the wider body of HNRT literature and knowledge regarding the optimal treatment margin and DIR-facilitated dose accumulation. *Sub-sections* 5.6.1 and 5.6.2 discuss treatment margin optimisation and DIR-facilitated dose accumulation, respectively, and are presented in the context of the research described in this dissertation and future developments.

5.6.1 Treatment margin optimisation

Chapter Four found that both 5 and 3 mm margin plans for VMAT HNRT are generally robust for anatomical changes which suggests the possibility of further PTV margin reduction. The retrospective DIR-facilitated dose accumulation studies by van Kranen et al. [142] and Wu et al [110] investigated the impact of anatomical changes to the dose delivery during modulated HNRT for 5, 3 and 0 mm PTV margin expansions. The study of Wu et al [110] found that the coverage of CTVs at the end of treatment was not affected by anatomical changes based on the ratio of multiple DVH metrics relative to the corresponding planning metrics for all PTV margin expansions. The study by van Kranen et al. [142] found that 27% (20/73) of CTVs were underdosed ($D_{99\%}^{acc}$ less than 95% of D^{presc}) for a 0 mm PTV margin expansion while for a 3 mm expansion, this occurred in only 3% (2/73) of CTVs. Based on the findings of the two preceding studies [110,142] and the results in *Chapter Four*, it seems likely that the PTV margin may be further optimisable to an expansion less than 3 mm. However, it is worth reiterating that a direct comparison of the results in these three studies is not possible due to differing analyses that were used to quantify the CTV coverage. Furthermore, all three studies mentioned in this paragraph so far evaluated treatment plan robustness in the presence of anatomical changes only and did not consider other treatment related uncertainties such as the geometrical accuracy of the treatment machine and the geometrical uncertainty of the target volume definition. Reducing the PTV margin based on the improvement of one aspect of the treatment, such as the robustness of plans for anatomical changes, could still be dangerous and may lead to CTV underdosage if the additional treatment related uncertainties are not considered.

In addition to the lack of consensus on the optimal PTV margin expansion in HNRT, a lack of consensus on the optimal CTV margin also exists [1–4,32–34,171,193,218,240,241,247]. Both of these ambiguities may at least partially be due to the uncertainty in GTV definition, which is seldom quantified. Unfortunately, studies using histopathological validation (the gold standard) [37,38,214] to assess the target volume definition uncertainty are limited, owing to their complexity. The implementation of a procedure that facilitates peer review of target volume delineations [248,249] could

reduce the uncertainty in target definition without considerably increasing physicians' workloads. A further reduction to the target volume definition uncertainty may be achieved with the routine inclusion of high-contrast MRI scans [250] and PET [251] images in the contouring process as they are generally considered to reduce the variation [252]. Although it should be noted that geometric distortions [253] and non-standardised imaging sequences [254] can restrict the effective use of MRI for target delineation. A current recommendation is that GTV delineation should be conducted on the pCT rigidly registered to an MRI scan that is also acquired in the treatment position [255,256]. However the availability of a diagnostic position MRI scan, not to mention a treatment position MRI scan, is not standardly available in all centres. Encouragingly, a recent study by Chuter et al. [257] found that the deformable registration of MRI scan acquired in the diagnostic position to the pCT facilitated a significant improvement in GTV delineation accuracy compared to contouring solely on the pCT. This finding could strengthen the case for departments to routinely include MRI scans. The use of artificial intelligence (AI) for structure definition [258], and in radiation oncology in general [259,260] is currently a “hot topic”. While promising methods such as those developed by Huang et al. [258] have demonstrated automatic AI delineation of the GTV with high accuracy and efficiency, AI methods for structure definition in general differ in performance [261] and can struggle in regions of homogeneous image contrast [262]. In these cases, incomplete and inaccurate definitions can be generated which require laborious manual correction. Nevertheless, the use of AI for structure definition seems to be a promising tool to assist the clinician in HN cancer management in the future [263,264].

Improvements in the accuracy of target volume definition may lead to more favourable patient outcomes but a definitive proof of safe PTV margin reduction is likely to only be achieved by reporting clinical results in conjunction with a comprehensive description of all relevant treatment details. For example, plan design, constraint adoption, prioritisation of OARs and target coverage, technical aspects of treatment accuracy, robustness of the planning solution for anatomical changes and patient positioning accuracy are aspects of treatment which can all influence treatment outcomes [5] irrespective of the accuracy of target volume definition. Clinical results reports should also standardly include the locations of recurrences and use consistent definitions for recurrence locations to increase insight of how

treatment margins can be optimised. In literature, a "marginal recurrence" has been defined by 20% to 95% of the LRR volume falling within the planned dose 95% isodose line [265] but also if 20% to 95% of the LRR volume was within the CTV [266]. It is recommended that future studies reporting on marginal recurrences use isodose lines of the DIR-facilitated accumulated dose distribution to define LRR locations. In addition, the actually delivered doses to the recurrence volumes as assessed by dose accumulation should be recorded which may increase insight into the variation of patient-specific tumour radio-sensitivity [267,268].

5.6.2 DIR-facilitated dose accumulation

In *Chapter Three* the intended clinical workflow based on DIR-facilitated dose accumulation to assess the actually delivered dose was defined and was subsequently found to be sufficiently accurate to assess the adequacy of treatment. This finding adds to a growing body of literature showing that DIR-facilitated dose accumulation is a tool that can be used to quantify the delivered dose [166,204,269]. The novel quantification technique developed in *Chapter Three* may draw greater attention to the accuracy and precision of DIR-facilitated dose accumulation, potentially progressing its clinical application and assisting to resolve the appropriateness of DIR-facilitated dose accumulation debate [139,143,166–169,236]. Considering the results in *Chapter Three* and the previously mentioned literature [166,204,269], DIR-facilitated dose accumulation can be utilised to objectively decide whether ART is necessary for individual patients. It is likely that a fully online ART procedure that includes online evaluation of the delivered dose (DIR-facilitated), re-planning [114,245] and subsequent delivery will be the standard of radiotherapy in the future. In fact, the integration of MRI with linear accelerators has already shown the feasibility of daily online re-planning [270] and improvements to the plan optimisation time [271] are expected to reduce the longer associated treatment times. An additional advantage to the online ART procedure is that a further reduction to the PTV margin can be achieved [272,273]. However, a number of limitations of DIR-facilitated dose accumulation still remain which are likely to restrict a large scale implementation of fully online ART supported with DIR-facilitated dose accumulation in the immediate future.

In general, DIR algorithms disregard changes occurring within target volumes and normal tissues at the cellular level, in particular for situations of mass and density alteration. Currently, the addition of local rigidity constraints [197,198] and constraints on inverse consistency and symmetry [147,148] more accurately inform the DVF regularisation but the inclusion of models describing tumour and normal tissue progression [274,275] is recommended for inclusion in future DIR algorithms. This limitation may itself represent an unacceptable level of risk to manufacturers. Furthermore, the technological demands of a fully online ART procedure are non-trivial in the context of a standard radiotherapy centre [114,139]. Another limitation of DIR-facilitated dose accumulation in an online ART procedure is that the accumulated dose will require immediate radiation oncologist review and approval. This requirement is likely to represent a prohibitively large increase in clinicians' workloads for large scale implementation. In the immediate future, studies (including *Chapter One*) that have attempted to predict patients for ART [79,92,104,127,207,276] will still be beneficial to gain understanding on where specific ART approaches are necessary or on when specific anatomical changes lead to a clinically irrelevant change in dosimetry.

As DIR-facilitated dose accumulation becomes more routine in the clinical setting, it must be realised that the validation of this method is not trivial and will increase the demand on department resources. The increase in complexity of the tools informing online ART must be met with comprehensive understanding of the limitations and clinical risk associated with these novel techniques. It is up to the discretion of the users to select an approach that works optimal for a specific application.

6 CONCLUSION

This dissertation aimed to enable a safe PTV margin reduction in HNRT by investigating the application of DIR in combination with EWMA SPC charts and DIR-facilitated dose accumulation to identify patients who may be at risk of unacceptable loss in target coverage and/or overdosage to normal tissues. EWMA SPC enabled detection of systematic anatomical changes and DIR-facilitated dose accumulation was sufficiently accurate to assess the adequacy of the delivered dose. The patient-specific nature of target coverage loss prevented the definition of generic guidelines for ART based on observed anatomical changes or patient positioning trends. However, the developed DIR-facilitated dose accumulation workflow will facilitate the creation of generic guidelines to objectively decide whether treatment adaptation is necessary for individual patients. Both 5 and 3 mm margin plans for VMAT HNRT were found to be robust for anatomical changes, while tight- or under-coverage of target volumes already present at the treatment planning stage turned out to be an important factor that may impact target coverage and potentially result in clinically unacceptable dose delivery. Assessment of the target coverage using DVH parameters only was not sufficient and patient-specific verification of the delivered dose using the methodology described in this dissertation is recommended.

REFERENCES

[1] Chen AM, Yu Y, Daly ME, Farwell GD, Benedict SH, Purdy JA. Long-term experience with reduced planning target volume margins and intensity-modulated radiotherapy with daily image-guidance for head and neck cancer. Head Neck 2014;36:1766–72. doi:https://doi.org/10.1002/hed.23532.

[2] Chen AM, Farwell DG, Luu Q, Donald PJ, Perks J, Purdy JA. Evaluation of the planning target volume in the treatment of head and neck cancer with intensity-modulated radiotherapy: What is the appropriate expansion margin in the setting of daily image guidance? Int J Radiat Oncol Biol Phys 2011;81:943–9. doi:10.1016/j.ijrobp.2010.07.017.

[3] Navran A, Heemsbergen W, Janssen T, Hamming-Vrieze O, Jonker M, Zuur C, et al. The impact of margin reduction on outcome and toxicity in head and neck cancer patients treated with image-guided volumetric modulated arc therapy (VMAT). Radiother Oncol 2018;130:25–31. doi:10.1016/j.radonc.2018.06.032.

[4] Franzese C, Fogliata A, Franceschini D, Navarria P, Cozzi L, Tomatis S, et al. Impact of hypofractionated schemes in radiotherapy for locally advanced head and neck cancer patients. Laryngoscope 2019:1–8. doi:10.1002/lary.28048.

[5] Das IJ, Cheng CW, Chopra KL, Mitra RK, Srivastava SP, Glatstein E. Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and treatment planning systems. J Natl Cancer Inst 2008;100:300–7. doi:10.1093/jnci/djn020.

[6] Argiris A, Karamouzis M V, Raben D, Ferris RL. Head and neck cancer. Lancet (London, England) 2008;371:1695–709. doi:10.1016/S0140-6736(08)60728-X.

[7] Argiris A, Eng C. Epidemiology, staging, and screening of head and neck cancer. Head neck cancer, Springer US; 2004, p. 15–60.

[8] Glastonbury CM, Harnsberger HR, Michel MA. Head and Neck Cancer: State of the Art Diagnosis, Staging, and Surveillance. Amirsys; 2012.

[9] Blot WJ, McLaughlin JK, Winn DM, Austin DF, Greenberg RS, Preston-Martin S, et al. Smoking and drinking in relation to oral and pharyngeal cancer. Cancer Res 1988;48:3282–7.

doi:10.1007/s10552-007-9026-4.

[10] Fakhry C, Gillison ML. Clinical implications of human papillomavirus in head and neck cancers. J Clin Oncol 2006;24:2606–11. doi:10.1200/JCO.2006.06.1291.

[11] Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol 2011;29:4294–301. doi:10.1200/JCO.2011.36.4596.

[12] Takes RP, Rinaldo A, Silver CE, Haigentz M, Woolgar JA, Triantafyllou A, et al. Distant metastases from head and neck squamous cell carcinoma. Part I. Basic aspects. Oral Oncol 2012;48:775–9. doi:10.1016/j.oraloncology.2012.03.013.

[13] Argiris A, Brockstein BE, Haraf DJ, Stenson KM, Mittal BB, Kies MS, et al. Competing causes of death and second primary tumors in patients with locoregionally advanced head and neck cancer treated with chemoradiotherapy. Clin Cancer Res 2004;10:1956–62. doi:10.1158/1078-0432.CCR-03-1077.

[14] Vokes EE, Weichselbaum RR, Lippman SM, Hong WK. Head and Neck Cancer. N Engl J Med 1993;328:184–94.

[15] Hall SF, Groome PA, Rothwell D. THE IMPACT OF COMORBIDITY ON THE SURVIVAL OF PATIENTS WITH SQUAMOUS CELL CARCINOMA OF THE HEAD AND NECK. Head Neck 2000;22:317–22.

[16] New Zealand Ministry of Health. Cancer: New registrations and deaths 2016. 2018.

[17] New Zealand Ministry of Health. Cancer Patient Survival. 2015.

[18] Bhide SA, Nutting CM. Advances in radiotherapy for head and neck cancer. Oral Oncol 2010;46:439–41. doi:10.1016/j.oraloncology.2010.03.005.

[19] Podgorsak EB. Radiation oncology physics. a handbook for teachers and students/EB Podgorsak. Vienna: International Atomic Energy Agency; 2005.

[20] Thoms J, Bristow RG. DNA repair targeting and radiotherapy: A focus on the therapeutic ratio. Semin Radiat Oncol 2010;20:217–22. doi:10.1016/j.semradonc.2010.06.003.

[21] Mendenhall WM, Amdur RJ, Palta JR. Intensity-modulated radiotherapy in the standard management of head and neck cancer: Promises and pitfalls. J Clin Oncol 2006;24:2618–23. doi:10.1200/JCO.2005.04.7225.

[22] Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. Med Phys 2008;35:310–7. doi:10.1118/1.2818738.

[23] Teoh M, Clark CH, Wood K, Whitaker S, Nisbet A. Volumetric modulated arc therapy: A

review of current literature and clinical use in practice. Br J Radiol 2011;84:967–96. doi:10.1259/bjr/22373346.

[24] Boyer AL, Geis P, Grant W, Carol M. Modulated beam conformal therapy for head and neck tumors. Int J Radiat Oncol Biol Phys 1997;39:227–36.

[25] Eisbruch A, Ten Haken RK, Kim HM, Marsh LH, Ship JA. Dose, volume, and function relationships in parotid salivary glands following conformal and intensity-modulated irradiation of head and neck cancer. Int J Radiat Oncol Biol Phys 1999;45:577–87. doi:10.1016/S0360-3016(99)00247-3.

[26] Ahnesjö A, Aspradakis MM. Dose calculations for external photon beams in radiotherapy. Phys Med Biol 1999;44. doi:10.1088/0031-9155/44/11/201.

[27] Webb S. The physics of conformal radiotherapy: advances in technology (PBK). 1st Editio. Boca Raton: CRC Press; 1997. doi:https://doi.org/10.1201/9780367806545.

[28] Holt A, Van Gestel D, Arends MP, Korevaar EW, Schuring D, Kunze-Busch MC, et al. Multi-institutional comparison of volumetric modulated arc therapy vs. intensity-modulated radiation therapy for head-and-neck cancer: A planning study. Radiat Oncol 2013;8:1–11. doi:10.1186/1748-717X-8-26.

[29] Verbakel W, Cuijpers JP, Hoffmans D, Bieker M, Slotman BJ, Senan S. Volumetric Intensity-Modulated Arc Therapy Vs. Conventional IMRT in Head-and-Neck Cancer: A Comparative Planning and Dosimetric Study. Int J Radiat Oncol Biol Phys 2009;74:252–9. doi:10.1016/j.ijrobp.2008.12.033.

[30] Chen AM, Jennelle RLS, Grady V, Tovar A, Bowen K, Simonin P, et al. Prospective Study of Psychosocial Distress Among Patients Undergoing Radiotherapy for Head and Neck Cancer. Int J Radiat Oncol 2009;73:187–93. doi:10.1016/j.ijrobp.2008.04.010.

[31] ICRU. ICRU Report 50 - Prescribing, Recording, and Reporting Photon Beam Therapy. 1993.

[32] Hansen CR, Johansen J, Samsøe E, Andersen E, Petersen JBB, Jensen K, et al. Consequences of introducing geometric GTV to CTV margin expansion in DAHANCA contouring guidelines for head and neck radiotherapy. Radiother Oncol 2018;126:43–7. doi:10.1016/j.radonc.2017.09.019.

[33] Zukauskaite R, Hansen CR, Grau C, Samsøe E, Johansen J, Petersen JBB, et al. Local recurrences after curative IMRT for HNSCC: Effect of different GTV to high-dose CTV margins. Radiother Oncol 2018;126:48–55. doi:10.1016/j.radonc.2017.11.024.

[34] Grégoire V, Evans M, Le QT, Bourhis J, Budach V, Chen A, et al. Delineation of the primary tumour Clinical Target Volumes (CTV-P) in laryngeal, hypopharyngeal, oropharyngeal and

oral cavity squamous cell carcinoma: AIRO, CACA, DAHANCA, EORTC, GEORCC, GORTEC, HKNPCSG, HNCIG, IAG-KHT, LPRHHT, NCIC CTG, NCRI, NRG Oncolog. Radiother Oncol 2018;126:3–24. doi:10.1016/j.radonc.2017.10.016.

[35] Van Herk M. Errors and Margins in Radiotherapy. Semin Radiat Oncol 2004;14:52–64. doi:10.1053/j.semradonc.2003.10.003.

[36] Van Herk M, Remeijer P, Rasch C, Lebesque J V. The probability of correct target dosage: Dose-population histograms for deriving treatment margins in radiotherapy. Int J Radiat Oncol Biol Phys 2000;47:1121–35. doi:10.1016/S0360-3016(00)00518-6.

[37] Caldas-Magalhaes J, Kooij N, Ligtenberg H, Jager E a, Schakel T, Kasperts N, et al. The accuracy of target delineation in laryngeal and hypopharyngeal cancer. Acta Oncol (Madr) 2015:1–7. doi:10.3109/0284186X.2015.1006401.

[38] Daisne J-F, Duprez T, Weynand B, Lonneux M, Hamoir M, Reychler H, et al. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology 2004;233:93–100. doi:10.1148/radiol.2331030660.

[39] Peters LJ, Goepfert H, Ang KK, Byers RM, Maor MH, Guillamondegui O, et al. Evaluation of the dose for postoperative radiation therapy of head and neck cancer: First report of a prospective randomized trial n.d.

[40] Bernier J, Cooper JS, Pajak TF, Van Glabbeke M, Bourhis J, Forastiere A, et al. Defining risk levels in locally advanced head and neck cancers: A comparative analysis of concurrent postoperative radiation plus chemotherapy trials of the EORTC (#22931) and RTOG (#9501). Head Neck 2005;27:843–50. doi:10.1002/hed.20279.

[41] Langendijk JA, Slotman BJ, Van Der Waal I, Doornaert P, Berkof J, Leemans CR. Risk-group definition by recursive partitioning analysis of patients with squamous cell head and neck carcinoma treated with surgery and postoperative radiotherapy. Cancer 2005;104:1408–17. doi:10.1002/cncr.21340.

[42] International Commission on Radiation Units and Measurements. ICRU report 62: Prescribing, recording and reporting photon beam therapy (supplement to ICRU report 50). 1999.

[43] McKenzie A, van Herk M, Mijnheer BJ. Margins for geometric uncertainty around organs at risk in radiotherapy. Radiother Oncol 2002;62:299–307.

[44] Landberg T, Chavaudra J, Dobbs J, Gerard J-P, Hanks G, Horiot J-C, et al. ICRU Report 62. vol. os32. 1999. doi:10.1093/jicru/os32.1.Report62.

[45] Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of

Normal Tissue Complication Probability Models in the Clinic. Int J Radiat Oncol Biol Phys 2010;76. doi:10.1016/j.ijrobp.2009.07.1754.

[46] Drzymala RE, Mohan R, Brewster L, Chu J, Goitein M, Harms W, et al. Dose-volume histograms. Int J Radiat Oncol Biol Phys 1991;21:71–8. doi:10.1016/0360-3016(91)90168-4.

[47] Doornaert P, Verbakel WFAR, Bieker M, Slotman BJ, Senan S. RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy. Int J Radiat Oncol Biol Phys 2011;79:429–35. doi:10.1016/j.ijrobp.2009.11.014.

[48] Hall EJ, Wuu CS. Radiation-induced second cancers: The impact of 3D-CRT and IMRT. Int J Radiat Oncol Biol Phys 2003;56:83–8. doi:10.1016/S0360-3016(03)00073-7.

[49] Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. J Clin Oncol 2008;26:3770–6. doi:10.1200/JCO.2007.14.6647.

[50] Epstein JB, Emerton S, Kolbinson DA, Le ND, Phillips N, Stevenson-Moore P, et al. Quality of life and oral function following radiotherapy for head and neck cancer. Head Neck 1999;21:1–11. doi:10.1002/(SICI)1097-0347(199901)21:1<1::AID-HED1>3.0.CO;2-4.

[51] Bjordal K, Kaasa S, Mastekaasa A. Quality of life in patients treated for head and neck cancer: A follow-up study 7 to 11 years after radiotherapy. Int J Radiat Oncol 1994;28:847–56. doi:10.1016/0360-3016(94)90104-X.

[52] Chen AM, Farwell DG, Luu Q, Vazquez EG, Lau DH, Purdy JA. Intensity-modulated radiotherapy is associated with improved global quality of life among long-term survivors of head-and-neck cancer. Int J Radiat Oncol Biol Phys 2012;84:170–5. doi:10.1016/j.ijrobp.2011.11.026.

[53] Gupta T, Agarwal J, Jain S, Phurailatpam R, Kannan S, Ghosh-Laskar S, et al. Three-dimensional conformal radiotherapy (3D-CRT) versus intensity modulated radiation therapy (IMRT) in squamous cell carcinoma of the head and neck: A randomized controlled trial. Radiother Oncol 2012;104:343–8. doi:10.1016/j.radonc.2012.07.001.

[54] Jensen SB, Pedersen AML, Vissink A, Andersen E, Brown CG, Davies AN, et al. A systematic review of salivary gland hypofunction and xerostomia induced by cancer therapies: Prevalence, severity and impact on quality of life. Support Care Cancer 2010;18:1039–60. doi:10.1007/s00520-010-0827-8.

[55] Dijkema T, Raaijmakers CPJ, Braam PM, Roesink JM, Monninkhof EM, Terhaard CHJ.

Xerostomia: A day and night difference. Radiother Oncol 2012;104:219–23. doi:10.1016/j.radonc.2012.06.004.

[56] Deasy JO, Moiseenko V, Marks L, Chao KSC, Nam J, Eisbruch A. Radiotherapy Dose-Volume Effects on Salivary Gland Function. Int J Radiat Oncol Biol Phys 2010;76:58–63. doi:10.1016/j.ijrobp.2009.06.090.

[57] Lee C, Langen KM, Lu W, Haimerl J, Schnarr E, Ruchala KJ, et al. Assessment of Parotid Gland Dose Changes During Head and Neck Cancer Radiotherapy Using Daily Megavoltage Computed Tomography and Deformable Image Registration. Int J Radiat Oncol Biol Phys 2008;71:1563–71. doi:10.1016/j.ijrobp.2008.04.013.

[58] O'Daniel JC, Garden AS, Schwartz DL, Wang H, Ang KK, Ahamad A, et al. Parotid Gland Dose in Intensity-Modulated Radiotherapy for Head and Neck Cancer: Is What You Plan What You Get? Int J Radiat Oncol Biol Phys 2007;69:1290–6. doi:10.1016/j.ijrobp.2007.07.2345.

[59] Dijkema T, Raaijmakers CPJ, Ten Haken RK, Roesink JM, Braam PM, Houweling AC, et al. Parotid gland function after radiotherapy: The combined Michigan and Utrecht experience. Int J Radiat Oncol Biol Phys 2010;78:449–53. doi:10.1016/j.ijrobp.2009.07.1708.

[60] Hunter KU, Fernandes LL, Vineberg KA, McShan D, Antonuk AE, Cornwall C, et al. Parotid glands dose-effect relationships based on their actually delivered doses: Implications for adaptive replanning in radiation therapy of head-and-neck cancer. Int J Radiat Oncol Biol Phys 2013;87:676–82. doi:10.1016/j.ijrobp.2013.07.040.

[61] Pow EHN, Kwong DLW, McMillan AS, Wong MCM, Sham JST, Leung LHT, et al. Xerostomia and quality of life after intensity-modulated radiotherapy vs. conventional radiotherapy for early-stage nasopharyngeal carcinoma: Initial report on a randomized controlled clinical trial. Int J Radiat Oncol Biol Phys 2006;66:981–91. doi:10.1016/j.ijrobp.2006.06.013.

[62] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): A phase 3 multicentre randomised controlled trial. Lancet Oncol 2011;12:127–36. doi:10.1016/S1470-2045(10)70290-4.

[63] Caudell JJ, Schaner PE, Meredith RF, Locher JL, Nabell LM, Carroll WR, et al. Factors Associated With Long-Term Dysphagia After Definitive Radiotherapy for Locally Advanced Head-and-Neck Cancer. Int J Radiat Oncol Biol Phys 2009;73:410–5. doi:10.1016/j.ijrobp.2008.04.048.

[64] Russi EG, Corvò R, Merlotti A, Alterio D, Franco P, Pergolizzi S, et al. Swallowing

dysfunction in head and neck cancer patients treated by radiotherapy: Review and recommendations of the supportive task group of the Italian Association of Radiation Oncology. Cancer Treat Rev 2012;38:1033–49. doi:10.1016/j.ctrv.2012.04.002.

[65] Schindler A, Denaro N, Russi EG, Pizzorni N, Bossi P, Merlotti A, et al. Dysphagia in head and neck cancer patients treated with radiotherapy and systemic therapies: Literature review and consensus. Crit Rev Oncol Hematol 2015;96:372–84. doi:10.1016/j.critrevonc.2015.06.005.

[66] Fua TF, Corry J, Milner AD, Cramb J, Walsham SF, Peters LJ. Intensity-modulated radiotherapy for nasopharyngeal carcinoma: Clinical correlation of dose to the pharyngo-esophageal axis and dysphagia. Int J Radiat Oncol Biol Phys 2007;67:976–81. doi:10.1016/j.ijrobp.2006.10.028.

[67] Feng FY, Kim HM, Lyden TH, Haxer MJ, Feng M, Worden FP, et al. Intensity-Modulated Radiotherapy of Head and Neck Cancer Aiming to Reduce Dysphagia: Early Dose-Effect Relationships for the Swallowing Structures. Int J Radiat Oncol Biol Phys 2007;68:1289–98. doi:10.1016/j.ijrobp.2007.02.049.

[68] Eisbruch A, Schwartz M, Rasch C, Vineberg K, Damen E, Van As CJ, et al. Dysphagia and aspiration after chemoradiotherapy for head-and-neck cancer: Which anatomic structures are affected and can they be spared by IMRT? Int J Radiat Oncol Biol Phys 2004;60:1425–39. doi:10.1016/j.ijrobp.2004.05.050.

[69] Friedman RB. Osteoradionecrosis: causes and prevention. NCI Monogr a Publ Natl Cancer Inst 1990:145–9.

[70] Reuther T, Schuster T, Mende U, Kübler AC. Osteoradionecrosis of the jaws as a side effect of radiotherapy of head and neck tumour patients - A report of a thirty year retrospective review. Int J Oral Maxillofac Surg 2003;32:289–95. doi:10.1054/ijom.2002.0332.

[71] Støre G, Boysen M. Mandibular osteoradionecrosis: Clinical behaviour and diagnostic aspects. Clin Otolaryngol Allied Sci 2000;25:378–84. doi:10.1046/j.1365-2273.2000.00367.x.

[72] Bhide SA, Harrington KJ, Nutting CM. Otological Toxicity After Postoperative Radiotherapy for Parotid Tumours. Clin Oncol 2007;19:77–82. doi:10.1016/j.clon.2006.11.007.

[73] Raaijmakers E, Engelen AM. Is sensorineural hearing loss a possible side effect of nasopharyngeal and parotid irradiation? A systematic review of the literature. Radiother Oncol 2002;65:1–7. doi:10.1016/S0167-8140(02)00211-6.

[74] Louise Kent M, Brennan MT, Noll JL, Fox PC, Burri SH, Hunter JC, et al. Radiation-induced trismus in head and neck cancer patients. Support Care Cancer 2008;16:305–9.

doi:10.1007/s00520-007-0345-5.

[75] Hall EJ, Giaccia AJ. Radiobiology for the Radiologist. 8th ed. Wolters Kluwer Health, 2018; 2006.

[76] Lacas B, Bourhis J, Overgaard J, Zhang Q, Grégoire V, Nankivell M, et al. Role of radiotherapy fractionation in head and neck cancers (MARCH): an updated meta-analysis. Lancet Oncol 2017;18:1221–37. doi:10.1016/S1470-2045(17)30458-8.

[77] Baskar R, Lee KA, Yeo R, Yeoh KW. Cancer and radiation therapy: Current advances and future directions. Int J Med Sci 2012;9:193–9. doi:10.7150/ijms.3635.

[78] Barker JL, Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys 2004;59:960–70. doi:10.1016/j.ijrobp.2003.12.024.

[79] Brouwer CL, Steenbakkers RJHM, Langendijk JA, Sijtsema NM. Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help? Radiother Oncol 2015;115:285–94. doi:10.1016/j.radonc.2015.05.018.

[80] Jaffray DA, Lindsay PE, Brock KK, Deasy JO, Tomé WA. Accurate Accumulation of Dose for Improved Understanding of Radiation Effects in Normal Tissue. Int J Radiat Oncol Biol Phys 2010;76:135–9. doi:10.1016/j.ijrobp.2009.06.093.

[81] Graff P, Kirby N, Weinberg V, Chen J, Yom SS, Lambert L, et al. The residual setup errors of different igrt alignment procedures for head and neck IMRT and the resulting dosimetric impact. Int J Radiat Oncol Biol Phys 2013;86:170–6. doi:10.1016/j.ijrobp.2012.10.040.

[82] Polat B, Wilbert J, Baier K, Flentje M, Guckenberger M. Nonrigid patient setup errors in the head-and-neck region. Strahlentherapie Und Onkol 2007;183:506–11. doi:10.1007/s00066-007-1747-5.

[83] van Kranen S, van Beek S, Rasch C, van Herk M, Sonke JJ. Setup Uncertainties of Anatomical Sub-Regions in Head-and-Neck Cancer Patients After Offline CBCT Guidance. Int J Radiat Oncol Biol Phys 2009;73:1566–73. doi:10.1016/j.ijrobp.2008.11.035.

[84] Zhang L, Garden AS, Lo J, Ang KK, Ahamad A, Morrison WH, et al. Multiple regions-of-interest analysis of setup uncertainties for head-and-neck cancer radiotherapy. Int J Radiat Oncol Biol Phys 2006;64:1559–69. doi:10.1016/j.ijrobp.2005.12.023.

[85] Moore SJ, Herst PM, Louwe RJW. Review of the patient positioning reproducibility in head-and-neck radiotherapy using Statistical Process Control. Radiother Oncol 2018;127:183–9.

doi:10.1016/j.radonc.2018.01.006.

[86] van Herk M. Different Styles of Image-Guided Radiotherapy. Semin Radiat Oncol 2007;17:258–67. doi:10.1016/j.semradonc.2007.07.003.

[87] Remeijer P, Geerlof E, Ploeger L, Gilhuijs K, Van Herk M, Lebesque J V. 3-D portal image analysis in clinical practice: An evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients. Int J Radiat Oncol Biol Phys 2000;46:1281–90. doi:10.1016/S0360-3016(99)00468-X.

[88] Moore SJ, Louwe RJW. Review of a multidisciplinary team approach to patient positioning in head-and-neck cancer: a quantitative analysis. A thesis submitted for the degree of Bachelor of Radiation Therapy with Honours at the University of Otago, Dunedin, New Zealand, 2014.

[89] Lees J. Incidence of weight loss in head and neck cancer patients on commencing radiotherapy treatment at a regional oncology centre. Eur J Cancer Care (Engl) 1999;8:133–6. doi:10.1046/j.1365-2354.1999.00156.x.

[90] Munshi A, Pandey MB, Durga T, Pandey KC, Bahadur S, Mohanti BK. Weight Loss During Radiotherapy for Head and Neck Malignancies: What Factors Impact It? Nutr Cancer 2003;47:136–40. doi:http://dx.doi.org/10.1207/s15327914nc4702_5.

[91] Bhide SA, Davies M, Burke K, McNair HA, Hansen V, Barbachano Y, et al. Weekly Volume and Dosimetric Changes During Chemoradiotherapy With Intensity-Modulated Radiation Therapy for Head and Neck Cancer: A Prospective Observational Study. Int J Radiat Oncol Biol Phys 2010;76:1360–8. doi:10.1016/j.ijrobp.2009.04.005.

[92] Yan DF, Yan SX, Wang QD, Liao XB, Lu ZJ, Wang YX. Predictors for replanning in loco-regionally advanced nasopharyngeal carcinoma patients undergoing intensity-modulated radiation therapy: A prospective observational study. BMC Cancer 2013;13. doi:10.1186/1471-2407-13-548.

[93] Ahn PH, Chen C-CC, Ahn AI, Hong L, Scripes PG, Shen J, et al. Adaptive planning in intensity-modulated radiation therapy for head and neck cancers: Single-institution experience and clinical implications. Int J Radiat Oncol Biol Phys 2011;80:677–85. doi:10.1016/j.ijrobp.2010.03.014.

[94] Broggi S, Fiorino C, Dell'Oca I, Dinapoli N, Paiusco M, Muraglia A, et al. A two-variable linear model of parotid shrinkage during IMRT for head and neck cancer. Radiother Oncol 2010;94:206–12. doi:10.1016/j.radonc.2009.12.014.

[95] Wang X, Lu J, Xiong X, Zhu G, Ying H, He S, et al. Anatomic and dosimetric changes during the treatment course of intensity-modulated radiotherapy for locally advanced nasopharyngeal

carcinoma. Med Dosim 2010;35:151–7. doi:10.1016/j.meddos.2009.06.007.

[96] Giske K, Stoiber EM, Schwarz M, Stoll A, Muenter MW, Timke C, et al. Local setup errors in image-guided radiotherapy for head and neck cancer patients immobilized with a custom-made device. Int J Radiat Oncol Biol Phys 2011;80:582–9. doi:10.1016/j.ijrobp.2010.07.1980.

[97] Jin X, Hu W, Shang H, Yi J, Zhou Y, Xie C. CBCT-based volumetric and dosimetric variation evaluation of volumetric modulated arc radiotherapy in the treatment of nasopharyngeal cancer patients. Radiat Oncol 2013;8:279. doi:10.1186/1748-717x-8-279.

[98] Ricchetti F, Wu B, McNutt T, Wong J, Forastiere A, Marur S, et al. Volumetric change of selected organs at risk during IMRT for oropharyngeal cancer. Int J Radiat Oncol Biol Phys 2011;80:161–8. doi:10.1016/j.ijrobp.2010.01.071.

[99] Cozzolino M, Fiorentino A, Oliviero C, Pedicini P, Clemente S, Califano G, et al. Volumetric and Dosimetric Assessment by Cone-Beam Computed Tomography Scans in Head and Neck Radiation Therapy: A Monitoring in Four Phases of Treatment. Technol Cancer Res Treat 2013;13:325–35. doi:10.7785/tcrt.2012.500380.

[100] Popovtzer A, Cao Y, Feng FY, Eisbruch A. Anatomical changes in the pharyngeal constrictors after chemo-irradiation of head and neck cancer and their dose-effect relationships: MRI-based study. Radiother Oncol 2009;93:510–5. doi:10.1016/j.radonc.2009.05.013.

[101] Rathod S, Gupta T, Ghosh-Laskar S, Murthy V, Budrukkar A, Agarwal J. Quality-of-life (QOL) outcomes in patients with head and neck squamous cell carcinoma (HNSCC) treated with intensity-modulated radiation therapy (IMRT) compared to three-dimensional conformal radiotherapy (3D-CRT): Evidence from a prospective randomized s. Oral Oncol 2013;49:634–42. doi:10.1016/j.oraloncology.2013.02.013.

[102] Tribius S, Bergelt C. Intensity-modulated radiotherapy versus conventional and 3D conformal radiotherapy in patients with head and neck cancer: Is there a worthwhile quality of life gain? Cancer Treat Rev 2011;37:511–9. doi:10.1016/j.ctrv.2011.01.004.

[103] Jensen AD, Nill S, Huber PE, Bendl R, Debus J, Münter MW. A clinical concept for interfractional adaptive radiation therapy in the treatment of head and neck cancer. Int J Radiat Oncol Biol Phys 2012;82:590–6. doi:10.1016/j.ijrobp.2010.10.072.

[104] Hermans BCM, Persoon LCGG, Podesta M, Hoebers FJP, Verhaegen F, Troost EGC. Weekly kilovoltage cone-beam computed tomography for detection of dose discrepancies during (chemo)radiotherapy for head and neck cancer. Acta Oncol (Madr) 2015;54:1483–9. doi:10.3109/0284186X.2015.1061210.

[105] Lee C, Langen KM, Lu W, Haimerl J, Schnarr E, Ruchala KJ, et al. Evaluation of geometric

changes of parotid glands during head and neck cancer radiotherapy using daily MVCT and automatic deformable registration. Radiother Oncol 2008;89:81–8. doi:10.1016/j.radonc.2008.07.006.

[106] Hansen EK, Bucci MK, Quivey JM, Weinberg V, Xia P. Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2006;64:355–62. doi:10.1016/j.ijrobp.2005.07.957.

[107] Schwartz DL, Garden AS, Shah SJ, Chronowski G, Sejpal S, Rosenthal DI, et al. Adaptive radiotherapy for head and neck cancer—Dosimetric results from a prospective clinical trial. Radiother Oncol 2013;106:80–4. doi:10.1016/j.radonc.2012.10.010.

[108] Nishi T, Nishimura Y, Shibata T, Tamura M, Nishigaito N, Okumura M. Volume and dosimetric changes and initial clinical experience of a two-step adaptive intensity modulated radiation therapy (IMRT) scheme for head and neck cancer. Radiother Oncol 2013;106:85–9. doi:10.1016/j.radonc.2012.11.005.

[109] Castadot P, Geets X, Lee JA, Grégoire V. Adaptive functional image-guided IMRT in pharyngo-laryngeal squamous cell carcinoma: Is the gain in dose distribution worth the effort? Radiother Oncol 2011;101:343–50. doi:10.1016/j.radonc.2011.06.011.

[110] Wu Q, Chi Y, Chen PY, Krauss DJ, Yan D, Martinez A. Adaptive Replanning Strategies Accounting for Shrinkage in Head and Neck IMRT. Int J Radiat Oncol Biol Phys 2009;75:924–32. doi:10.1016/j.ijrobp.2009.04.047.

[111] Chen C, Lin X, Pan J, Fei Z, Chen L, Bai P. Is it necessary to repeat CT imaging and replanning during the course of intensity-modulated radiation therapy for locoregionally advanced nasopharyngeal carcinoma? Jpn J Radiol 2013;31:593–9. doi:10.1007/s11604-013-0225-5.

[112] Zhao L, Wan Q, Zhou Y, Deng X, Xie C, Wu S. The role of replanning in fractionated intensity modulated radiotherapy for nasopharyngeal carcinoma. Radiother Oncol 2011;98:23–7. doi:10.1016/j.radonc.2010.10.009.

[113] Keall P, Poulsen P, Booth JT. See, Think, and Act: Real-Time Adaptive Radiotherapy. Semin Radiat Oncol 2019;29:228–35. doi:10.1016/j.semradonc.2019.02.005.

[114] Sonke JJ, Aznar M, Rasch C. Adaptive Radiotherapy for Anatomical Changes. Semin Radiat Oncol 2019;29:245–57. doi:10.1016/j.semradonc.2019.02.007.

[115] Castadot P, Lee JA, Geets X, Gregoire V. Adaptive radiotherapy of head and neck cancer. YSRAO 2010;20:84–93. doi:10.1016/j.semradonc.2009.11.002.

[116] Yan D. Adaptive Radiotherapy: Merging Principle Into Clinical Practice. Semin Radiat Oncol

2010;20:79–83. doi:10.1016/j.semradonc.2009.11.001.

[117] Yan D, Frank V, Wong J, Martinez A. Adaptive radiation therapy. Phys Med Biol 1997;42:123–32.

[118] Fowler J. The linear-quadratic formula and progress in fractionated radiotherapy. Br J Radiol 1989;62:679–94.

[119] Chen AM, Daly ME, Cui J, Mathai M, Benedict SH, Purdy JA. Clinical outcomes among patients with head and neck cancer treated by intensity-modulated radiotherapy with and without adaptive replanning. Head Neck 2014;36:1541–6. doi:10.1002/hed.

[120] Schwartz DL, Garden AS, Thomas J, Chen Y, Zhang Y, Lewin J, et al. Adaptive radiotherapy for head-and-neck cancer: Initial clinical outcomes from a prospective trial. Int J Radiat Oncol Biol Phys 2012;83:986–93. doi:10.1016/j.ijrobp.2011.08.017.

[121] Yang H, Hu W, Wang W, Chen P, Ding W, Luo W. Replanning during intensity modulated radiation therapy improved quality of life in patients with nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys 2013;85:e47–54. doi:10.1016/j.ijrobp.2012.09.033.

[122] Van Kranen S, Mencarelli A, Van Beek S, Rasch C, Van Herk M, Sonke JJ. Adaptive radiotherapy with an average anatomy model: Evaluation and quantification of residual deformations in head and neck cancer patients. Radiother Oncol 2013;109:463–8. doi:10.1016/j.radonc.2013.08.007.

[123] Nuver TT, Hoogeman MS, Remeijer P, van Herk M, Lebesque J V. An Adaptive Off-Line Procedure for Radiotherapy of Prostate Cancer. Int J Radiat Oncol Biol Phys 2007;67:1559–67. doi:10.1016/j.ijrobp.2006.12.010.

[124] Bondar ML, Hoogeman MS, Mens JW, Quint S, Ahmad R, Dhawtal G, et al. Individualized nonadaptive and online-adaptive intensity-modulated radiotherapy treatment strategies for cervical cancer patients based on pretreatment acquired variable bladder filling computed tomography scans. Int J Radiat Oncol Biol Phys 2012;83:1617–23. doi:10.1016/j.ijrobp.2011.10.011.

[125] Wright P, Redpath AT, Høyer M, Muren LP. A method to individualize adaptive planning target volumes for deformable targets. Phys Med Biol 2009;54:7121–33. doi:10.1088/0031-9155/54/23/006.

[126] Lutkenhaus LJ, Visser J, De Jong R, Hulshof MCCM, Bel A. Evaluation of delivered dose for a clinical daily adaptive plan selection strategy for bladder cancer radiotherapy. Radiother Oncol 2015;116:51–6. doi:10.1016/j.radonc.2015.06.003.

[127] Huang H, Lu H, Feng G, Jiang H, Chen J, Cheng J, et al. Determining appropriate timing of

adaptive radiation therapy for nasopharyngeal carcinoma during intensity-modulated radiation therapy. Radiat Oncol 2015;10:1–9. doi:10.1186/s13014-015-0498-1.

[128] Ahunbay EE, Peng C, Chen GP, Narayanan S, Yu C, Lawton C, et al. An on-line replanning scheme for interfractional variations. Med Phys 2008;35:3607–15. doi:10.1118/1.2952443.

[129] Jaffray DA, Das S, Jacobs PM, Jeraj R, Lambin P. How Advances in Imaging Will Affect Precision Radiation Oncology. Int J Radiat Oncol Biol Phys 2018;101:292–8. doi:10.1016/j.ijrobp.2018.01.047.

[130] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132: Report. Med Phys 2017;44:e43–76. doi:10.1002/mp.12256.

[131] Varadhan R, Karangelis G, Krishnan K. A framework for deformable image registration validation in. J Appl Clin Med Phys 2014;14:1–27.

[132] Kessler ML. Image registration and data fusion in radiation therapy. Br J Radiol 2006;79:99–108. doi:10.1259/bjr/70617164.

[133] Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal 1998;2:243–60.

[134] Wang H, Dong L, O'Daniel J, Mohan R, Garden AS, Kian Ang K, et al. Validation of an accelerated "demons" algorithm for deformable image registration in radiation therapy. Phys Med Biol 2005;50:2887–905. doi:10.1088/0031-9155/50/12/011.

[135] VARIAN Medical Systems. Registration, SmartAdapt and Contouring Reference Guide 2014:1–71.

[136] Zhang L, Wang Z, Shi C, Long T, Xu XG. The impact of robustness of deformable image registration on contour propagation and dose accumulation for head and neck adaptive radiotherapy. J Appl Clin Med Phys 2018;19:185–94. doi:10.1002/acm2.12361.

[137] Thor M, Petersen JBB, Bentzen L, Høyer M, Muren LP. Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer. Acta Oncol (Madr) 2011;50:918–25. doi:10.3109/0284186X.2011.577806.

[138] Kumarasiri A, Siddiqui F, Liu C, Yechieli R, Shah M, Pradhan D, et al. Deformable image registration based automatic CT-to-CT contour propagation for head and neck adaptive radiotherapy in the routine clinical setting. Med Phys 2014;41. doi:10.1118/1.4901409.

[139] Chetty IJ, Rosu-Bubulac M. Deformable Registration for Dose Accumulation. Semin Radiat Oncol 2019;29:198–208. doi:10.1016/j.semradonc.2019.02.002.

[140] Lu W, Olivera GH, Chen Q, Ruchala KJ, Haimerl J, Meeks SL, et al. Deformable registration

of the planning image (kVCT) and the daily images (MVCT) for adaptive radiation therapy. Phys Med Biol 2006;51:4357–74. doi:10.1088/0031-9155/51/17/015.

[141] Schaly B, Kempe JA, Bauman GS, Battista JJ, Van Dyk J. Tracking the dose distribution in radiation therapy by accounting for variable anatomy. Phys Med Biol 2004;49:791–805. doi:10.1088/0031-9155/49/5/010.

[142] van Kranen S, Hamming-Vrieze O, Wolf AAL, Damen E, van Herk M, Sonke JJ. Head and Neck Margin Reduction With Adaptive Radiation Therapy: Robustness of Treatment Plans Against Anatomy Changes. Int J Radiat Oncol Biol Phys 2016;96:653–60. doi:10.1016/j.ijrobp.2016.07.011.

[143] Schultheiss TE, Tome WA, Orton CG. Point/counterpoint: it is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy. Med Phys 2012;39:6531–3. doi:10.1118/1.4722968.

[144] Fotina I, Hopfgartner J, Stock M, Steininger T, Lütgendorf-Caucig C, Georg D. Feasibility of CBCT-based dose calculation: Comparative analysis of HU adjustment techniques. Radiother Oncol 2012;104:249–56. doi:10.1016/j.radonc.2012.06.007.

[145] Veiga C, Lourenço AM, Mouinuddin S, Van Herk M, Modat M, Ourselin S, et al. Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm. Med Phys 2015;42:760–9. doi:10.1118/1.4905050.

[146] Bender ET, Tomé WA. The utilization of consistency metrics for error analysis in deformable image registration. Phys Med Biol 2009;54:5561–77. doi:10.1088/0031-9155/54/18/014.

[147] Guetter C, Xue H, Chefd'hotel C, Guehring J. Efficient symmetric and inverse-consistent deformable registration through interleaved optimization. Proc - Int Symp Biomed Imaging 2011:590–3. doi:10.1109/ISBI.2011.5872476.

[148] Christensen GE, Song JH, Lu W, Naqa I El, Low DA. Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry. Med Phys 2007;34:2155–63. doi:10.1118/1.2731029.

[149] Yang D, Li H, Low DA, Deasy JO, Naqa I El. A fast inverse consistent deformable image registration method based on symmetric optical flow computation. Conf Rec - Asilomar Conf Signals, Syst Comput 2008:1931–5. doi:10.1109/ACSSC.2008.5074766.

[150] Ramadaan IS, Peick K, Hamilton DA, Evans J, Iupati D, Nicholson A, et al. Validation of Varian's SmartAdapt® deformable image registration algorithm for clinical application. Radiat Oncol 2015;10:1–9. doi:10.1186/s13014-015-0372-1.

[151] Hvid CA, Elstrøm U V., Jensen K, Alber M, Grau C. Accuracy of software-assisted contour propagation from planning CT to cone beam CT in head and neck radiotherapy. Acta Oncol (Madr) 2016;55:1324–30. doi:10.1080/0284186X.2016.1185149.

[152] Wang H, Garden AS, Zhang L, Wei X, Ahamad A, Kuban DA, et al. Performance Evaluation of Automatic Anatomy Segmentation Algorithm on Repeat or Four-Dimensional Computed Tomography Images Using Deformable Image Registration Method. Int J Radiat Oncol Biol Phys 2008;72:210–9. doi:10.1016/j.ijrobp.2008.05.008.

[153] Castadot P, Lee JA, Parraga A, Geets X, Macq B, Grégoire V. Comparison of 12 deformable registration strategies in adaptive radiation therapy for the treatment of head and neck tumors. Radiother Oncol 2008;89:1–12. doi:10.1016/j.radonc.2008.04.010.

[154] Qi XS, Santhanam A, Neylon J, Min Y, Armstrong T, Sheng K, et al. Near real-time assessment of anatomic and dosimetric variations for head and neck radiation therapy via graphics processing unit-based dose deformation framework. Int J Radiat Oncol Biol Phys 2015;92:415–22. doi:10.1016/j.ijrobp.2015.01.033.

[155] Chalana V, Kim Y. A Methodology for Evaluation of Boundary Detection Algorithms on Medical Images. IEEE Trans Med Imaging 1997;16:642–52.

[156] Dice L. Measures of the Amount of Ecologic Association Between Species. Ecology 1945;26:297–302.

[157] Vinod SK, Jameson MG, Min M, Holloway LC. Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies. Radiother Oncol 2016;121:169–79. doi:10.1016/j.radonc.2016.09.009.

[158] Fotina I, Lütgendorf-Caucig C, Stock M, Pötter R, Georg D. Critical discussion of evaluation parameters for inter-observer variability in target definition for radiation therapy. Strahlentherapie Und Onkol 2012;188:160–7. doi:10.1007/s00066-011-0027-6.

[159] Bender ET, Hardcastle N, Tomé WA. On the dosimetric effect and reduction of inverse consistency and transitivity errors in deformable image registration for dose accumulation. Med Phys 2012;39:272–80. doi:10.1118/1.3666948.

[160] Kashani R, Lam K, Litzenberg D, Balter J. Technical note: A deformable phantom for dynamic modeling in radiation therapy. Med Phys 2007;34:199–201. doi:10.1118/1.2400612.

[161] Pukala J, Meeks SL, Staton RJ, Bova FJ, Langen KM, Mañon RR. A virtual phantom library for the quantification of deformable image registration uncertainties in patients with cancers of the head and neck. Med Phys 2013;40:111703. doi:10.1118/1.4823467.

[162] Rigaud B, Simon A, Castelli J, Gobeli M, Ospina Arango JD, Cazoulat G, et al. Evaluation of

deformable image registration methods for dose monitoring in head and neck radiotherapy. Biomed Res Int 2015;2015. doi:10.1155/2015/726268.

[163] Zhong H, Weiss E, Siebers J V. Assessment of dose reconstruction errors in image-guided radiation therapy. Phys Med Biol 2008;53:719–36. doi:10.1088/0031-9155/53/3/013.

[164] Nassef M, Simon A, Cazoulat G, Duménil A, Blay C, Lafond C, et al. Quantification of dose uncertainties in cumulated dose estimation compared to planned dose in prostate IMRT. Radiother Oncol 2016;119:129–36. doi:10.1016/j.radonc.2016.03.007.

[165] Salguero FJ, Saleh-Sayah NK, Yan C, Siebers J V. Estimation of three-dimensional intrinsic dosimetric uncertainties resulting from using deformable image registration for dose mapping. Med Phys 2011;38:343–53. doi:10.1118/1.3528201.

[166] Yeo UJ, Taylor ML, Supple JR, Smith RL, Dunn L, Kron T, et al. Is it sensible to "deform" dose 3D experimental validation of dose-warping. Med Phys 2012;39:5065–72. doi:10.1118/1.4736534.

[167] Zhong H, Chetty IJ. Caution Must Be Exercised When Performing Deformable Dose Accumulation for Tumors Undergoing Mass Changes During Fractionated Radiation Therapy. Int J Radiat Oncol Biol Phys 2017;97:182–3. doi:10.1016/j.ijrobp.2016.09.012.

[168] Hugo GD, Dial C, Siebers J V. In Regard to Zhong and Chetty. Int J Radiat Oncol 2017;99:1308–10. doi:10.1016/j.ijrobp.2017.08.047.

[169] Taylor ML, Siva S, Franich RD, Yeo UJ, Pham D, Supple J, et al. Comment on "It is not appropriate to 'deform' dose along with deformable image registration in adaptive radiotherapy" [Med. Phys. 39, 6531-6533 (2012)]. Med Phys 2013;40:17101. doi:10.1118/1.4771962.

[170] Rasch C, Steenbakkers R, Van Herk M. Target definition in prostate, head, and neck. Semin Radiat Oncol 2005;15:136–45. doi:10.1016/j.semradonc.2005.01.005.

[171] Suzuki M, Nishimura Y, Nakamatsu K, Okumura M, Hashiba H, Koike R, et al. Analysis of interfractional set-up errors and intrafractional organ motions during IMRT for head and neck tumors to define an appropriate planning target volume (PTV)- and planning organs at risk volume (PRV)-margins. Radiother Oncol 2006;78:283–90. doi:10.1016/j.radonc.2006.03.006.

[172] Sanguineti G, Gunn GB, Endres EJ, Chaljub G, Cheruvu P, Parker B. Patterns of Locoregional Failure After Exclusive IMRT for Oropharyngeal Carcinoma. Int J Radiat Oncol Biol Phys 2008;72:737–46. doi:10.1016/j.ijrobp.2008.01.027.

[173] Schoenfeld GO, Amdur RJ, Morris CG, Li JG, Hinerman RW, Mendenhall WM. Patterns of Failure and Toxicity after Intensity-Modulated Radiotherapy for Head and Neck Cancer. Int J

Radiat Oncol Biol Phys 2008;71:377–85. doi:10.1016/j.ijrobp.2007.10.010.

[174] Orlandi E, Tomatis S, Potepan P, Bossi P, Mongioj V, Carrara M, et al. Critical analysis of locoregional failures following intensity- modulated radiotherapy for nasopharyngeal carcinoma. Futur Oncol 2013;9:103–14.

[175] Brouwer CL, Steenbakkers RJHM, Bourhis J, Budach W, Grau C, Grégoire V, et al. CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. Radiother Oncol 2015;117:83–90. doi:10.1016/j.radonc.2015.07.041.

[176] Grégoire V, Levendag P, Ang KK, Bernier J, Braaksma M, Budach V, et al. CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC, RTOG consensus guidelines. Radiother Oncol 2003;69:227–36. doi:10.1016/j.radonc.2003.09.011.

[177] Chao KSC, Bhide S, Chen H, Asper J, Bush S, Franklin G, et al. Reduce in Variation and Improve Efficiency of Target Volume Delineation by a Computer-Assisted System Using a Deformable Image Registration Approach. Int J Radiat Oncol Biol Phys 2007;68:1512–21. doi:10.1016/j.ijrobp.2007.04.037.

[178] Hardcastle N, Tomé WA, Cannon DM, Brouwer CL, Wittendorp PW, Dogan N, et al. A multi-institution evaluation of deformable image registration algorithms for automatic organ delineation in adaptive head and neck radiotherapy. Radiat Oncol 2012;7:90. doi:10.1186/1748-717X-7-90.

[179] Mencarelli A, van Beek S, van Kranen SR, Rasch C, van Herk M, Sonke JJ. Validation of deformable registration in head and neck cancer using analysis of variance. Med Phys 2012;39:6879–84. doi:10.1016/s0167-8140(12)72906-7.

[180] van Kranen S, van Beek S, Mencarelli A, Rasch C, van Herk M, Sonke JJ. Correction strategies to manage deformations in head-and-neck radiotherapy. Radiother Oncol 2010;94:199–205. doi:10.1016/j.radonc.2009.12.016.

[181] Yu Y, Michaud AL, Sreeraman R, Liu T, Purdy JA, Chen AM, et al. Comparison of daily versus nondaily image-guided radiotherapy protocols for patients treated with intensity-modulated radiotherapy for head and neck cancer. Head Neck 2014;36:992–7. doi:10.1002/hed.21421.

[182] Chen AM, Daly ME, Farwell G, Vazquez E, Courquin J, Lau DH, et al. Quality of life among long-term survivors of head and neck cancer treated by intensity-modulated radiotherapy. JAMA Otolaryngol - Head Neck Surg 2014;140:129–33. doi:10.1001/jamaoto.2013.5988.

[183] Franzese C, Fogliata A, Clerici E, Franceschini D, Villa E, D'Agostino G, et al. Toxicity profile and early clinical outcome for advanced head and neck cancer patients treated with simultaneous integrated boost and volumetric modulated arc therapy. Radiat Oncol 2015;10:1–10. doi:10.1186/s13014-015-0535-0.

[184] Merckel LG, Hamilton DA, Evans J, Kwon HJ, Louwe RJW. Results of radical (chemo) radiotherapy of oropharynx cancer using three-dimensional conformal radiation therapy or volumetric modulated arc therapy: a matched-pair analysis. Unpubl Dep Results 2018.

[185] Ayyalusamy A, Vellaiyan S, Shanmugam S, Ilamurugu A, Gandhi A, Shanmugam T, et al. Feasibility of offline head & neck adaptive radiotherapy using deformed planning CT electron density mapping on weekly cone beam computed tomography. Br J Radiol 2016;90:20160420.

[186] Eiland RB, Maare C, SjöStröm D, Samsøe E, Behrens CF. Dosimetric and geometric evaluation of the use of deformable image registration in adaptive intensity-modulated radiotherapy for head-and-neck cancer. J Radiat Res 2014;55:1002–8. doi:10.1093/jrr/rru044.

[187] Borror CM, Montgomery DC, Runger GC. Robustness of the EWMA Control Chart to Non-Normality. J Qual Technol 1999;31:309–16. doi:10.1080/00224065.1999.11979929.

[188] Pawlicki T, Whitaker M, Boyer AL. Statistical process control for radiotherapy quality assurance. Med Phys 2005;32:2777–86. doi:10.1118/1.2001209.

[189] Nordström F, Af Wetterstedt S, Johnsson S, Ceberg C, Bäck SÅJ. Control chart analysis of data from a multicenter monitor unit verification study. Radiother Oncol 2012;102:364–70. doi:10.1016/j.radonc.2011.11.016.

[190] Njeh CF, Parker BC. Evaluation of treatment plans using target and normal tissue DVHs is no longer appropriate. Med Phys 2015;42:2099–102. doi:10.1118/1.4903902.

[191] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): A phase 3 multicentre randomised controlled trial. Lancet Oncol 2011;12:127–36. doi:10.1016/S1470-2045(10)70290-4.

[192] Pacholke HD, Amdur RJ, Morris CG, Li JG, Dempsey JF, Hinerman RW, et al. Late xerostomia after intensity-modulated radiation therapy versus conventional radiotherapy. Am J Clin Oncol Cancer Clin Trials 2005;28:351–8. doi:10.1097/01.coc.0000158826.88179.75.

[193] Houghton F, Benson RJ, Tudor GSJ, Fairfoul J, Gemmill J, Dean JC, et al. An Assessment of Action Levels in Imaging Strategies in Head and Neck Cancer using TomoTherapy. Are Our Margins Adequate in the Absence of Image Guidance? Clin Oncol 2009;21:720–7. doi:10.1016/j.clon.2009.08.005.

[194] Yu Y, Michaud AL, Sreeraman R, Liu T, Purdy JA, Chen AM. Comparison of daily versus nondaily image-guided radiotherapy protocols for patients treated with intensity-modulated radiotherapy for head and neck cancer. Head Neck 2014;36:992–7. doi:10.1002/HED.

[195] Hunter JS. The Exponentially Weighted Moving Average. J Qual Technol 1986;18:203–10. doi:10.1080/00224065.1986.11979014.

[196] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, et al. Evaluation of measurement data—guide for the expression of uncertainty in measurement JCGM 100: 2008. Citado En Las 2008:167. doi:10.1373/clinchem.2003.030528.

[197] König L, Derksen A, Papenberg N, Haas B. Deformable image registration for adaptive radiotherapy with guaranteed local rigidity constraints. Radiat Oncol 2016;11:1–9. doi:10.1186/s13014-016-0697-4.

[198] Staring M, Klein S, Pluim JPW. A rigidity penalty term for nonrigid registration. Med Phys 2007;34:4098–108. doi:10.1118/1.2776236.

[199] Ciarmatori A, Maffei N, Mistretta GM, Ceroni P, Bernabei A, Meduri B, et al. Evaluation of the effectiveness of novel single-intervention adaptive radiotherapy strategies based on daily dose accumulation. Med Dosim 2019;44:379–84. doi:10.1016/j.meddos.2019.02.002.

[200] Kikinis R, Pieper SD, Vosburgh K. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. Intraoperative imaging image-guided Ther., New York, NY: Springer; 2014, p. 277–89.

[201] Pinter C, Lasso A, Wang A, Jaffray D, Fichtinger G. SlicerRT: Radiation therapy research toolkit for 3D Slicer. Med Phys 2012;39:6332–8. doi:10.1118/1.4754659.

[202] Sharp G, Li R, Wolfgang J, Chen G, Peroni M, Spadea MF, et al. Plastimatch: an open source software suite for radiotherapy image processing. Proc XVI'th Int Conf Use Comput Radiother 2010. doi:10.1016/j.autneu.2011.03.003.

[203] Guidi G, Maffei N, Vecchi C, Ciarmatori A, Mistretta GM, Gottardi G, et al. A support vector machine tool for adaptive tomotherapy treatments: Prediction of head and neck patients criticalities. Phys Medica 2015;31:442–51. doi:10.1016/j.ejmp.2015.04.009.

[204] Bohoudi O, Lagerwaard FJ, Bruynzeel AME, Niebuhr NI, Johnen W, Senan S, et al. End-to-end empirical validation of dose accumulation in MRI-guided adaptive radiotherapy for prostate cancer using an anthropomorphic deformable pelvis phantom. Radiother Oncol 2019. doi:10.1016/j.radonc.2019.09.014.

[205] VARIAN Medical Systems. TrueBeam 2.7 MR2 Customer Release Note 2017:1–56.

[206] Veiga C, McClelland J, Moinuddin S, Lourenço A, Ricketts K, Annkah J, et al. Toward

adaptive radiotherapy for head and neck patients: Feasibility study on using CT-to-CBCT deformable registration for “dose of the day” calculations. Med Phys 2014;41:6531–3. doi:10.1118/1.4864240.

[207] Lowther NJ, Hamilton DA, Kim H, Evans JM, Marsh SH, Louwe RJW. Monitoring anatomical changes of individual patients using statistical process control during head-and-neck radiotherapy. Phys Imaging Radiat Oncol 2019;9:21–7. doi:10.1016/j.phro.2018.12.004.

[208] Weistrand O, Svensson S. The ANACONDA algorithm for deformable image registration in radiotherapy. Med Phys 2015;42:40–53. doi:10.1118/1.4894702.

[209] Manning MA, Wu Q, Cardinale RM, Mohan R, Lauve AD, Kavanagh BD, et al. The effect of setup uncertainty on normal tissue sparing with IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2001;51:1400–9. doi:10.1016/S0360-3016(01)01740-0.

[210] Samuelsson A, Mercke C, Johansson K-A. Systematic set-up errors for IMRT in the head and neck region: effect on dose distribution. Radiother Oncol 2003;66:303–11.

[211] ICRU. ICRU Report 83 - Prescribing, Recording, and Reporting Photon Beam Therapy. vol. 10. 2010. doi:https://doi.org/10.1093/jicru/ndq002.

[212] Thomas SJ, Hoole ACF. The effect of optimization on surface dose in intensity modulated radiotherapy (IMRT). Phys Med Biol 2004;49:4919–28. doi:10.1088/0031-9155/49/21/005.

[213] Lowther NJ, Marsh SH, Louwe RJW. Quantifying the dose accumulation uncertainty after deformable image registration in head-and-neck radiotherapy. Radiother Oncol 2020;143:117–25. doi:https://doi.org/10.1016/j.radonc.2019.12.009.

[214] Campbell S, Poon I, Markel D, Vena D, Higgins K, Enepekides D, et al. Evaluation of microscopic disease in oral tongue cancer using whole-mount histopathologic techniques: Implications for the management of head-and-neck cancers. Int J Radiat Oncol Biol Phys 2012;82:574–81. doi:10.1016/j.ijrobp.2010.09.038.

[215] Apisarnthanarax S, Elliott DD, El-Naggar AK, Asper JA, Blanco A, Ang KK, et al. Determining optimal clinical target volume margins in head-and-neck cancer based on microscopic extracapsular extension of metastatic neck nodes. Int J Radiat Oncol Biol Phys 2006;64:678–83. doi:10.1016/j.ijrobp.2005.08.020.

[216] Brouwer CL, Steenbakkers RJHM, Bourhis J, Budach W, Grau C, Grégoire V, et al. CT-based delineation of organs at risk in the head and neck region Oncology and TROG consensus guidelines. Radiother Oncol 2015;117:1–291.

[217] Ooishi M, Motegi A, Kawashima M, Arahira S, Zenda S, Nakamura N, et al. Patterns of failure after postoperative intensity-modulated radiotherapy for locally advanced and recurrent

head and neck cancer. Jpn J Clin Oncol 2016;46:919–27. doi:10.1093/jjco/hyw095.

[218] Bayman E, Prestwich RJD, Speight R, Aspin L, Garratt L, Wilson S, et al. Patterns of Failure after Intensity-modulated Radiotherapy in Head and Neck Squamous Cell Carcinoma using Compartmental Clinical Target Volume Delineation. Clin Oncol 2014;26:636–42. doi:10.1016/j.clon.2014.05.001.

[219] Shakam A, Scrimger R, Liu D, Mohamed M, Parliament M, Field GC, et al. Dose-volume analysis of locoregional recurrences in head and neck IMRT, as determined by deformable registration: A prospective multi-institutional trial. Radiother Oncol 2011;99:101–7. doi:10.1016/j.radonc.2011.05.008.

[220] Kry S, Smith S, Weathers R, Stovall M. Skin Dose during Radiotherapy: A Summary and General Estimation Technique. J Appl Clin Med Phys 2012;13:20–34.

[221] Louwe RJW, Wendling M, Monshouwer R, Satherley T, Day RA, Greig L. Time-resolved dosimetry using a pinpoint ionization chamber as quality assurance for IMRT and VMAT. Med Phys 2015;42:1625–39. doi:10.1118/1.4914395.

[222] Purdy JA. Buildup/surface dose and exit dose measurements for a 6-MV linear accelerator. Med Phys 1986;13:259–62.

[223] Klein E, Purdy JA. Entrance and exit dose regions for a clinac-2100c. Int J Radiat Oncol Biol Phys 1993;27:429–35.

[224] Branchini M, Broggi S, Dell'Oca I, Cattaneo GM, Calandrino R, Di Muzio NG, et al. Skin dose calculation during radiotherapy of head and neck cancer using deformable image registration of planning and mega-voltage computed tomography scans. Phys Imaging Radiat Oncol 2018;8:44–50. doi:10.1016/j.phro.2018.11.008.

[225] Moore S, Louwe RJW. Review of patient positioning in head-and-neck radiotherapy using Statistical Process Control. 2017.

[226] Broggi S, Fiorino C, Dell'Oca I, Dinapoli N, Paiusco M, Muraglia A, et al. A two-variable linear model of parotid shrinkage during IMRT for head and neck cancer. Radiother Oncol 2010;94:206–12. doi:10.1016/j.radonc.2009.12.014.

[227] Montgomery DC. Introduction to Statistical Quality Control. 6th Edit. John Wiley & Sons; 2008.

[228] Murphy MJ, Salguero FJ, Siebers J V., Staub D, Vaman C. A method to estimate the effect of deformable image registration uncertainties on daily dose mapping. Med Phys 2012;39:573–80. doi:10.1118/1.3673772.

[229] Zhong H, Chetty IJ. Adaptive radiotherapy for NSCLC patients: Utilizing the principle of

energy conservation to evaluate dose mapping operations. Phys Med Biol 2017;62:4333–45. doi:10.1088/1361-6560/aa54a5.

[230] Hub M, Thieke C, Kessler ML, Karger CP. A stochastic approach to estimate the uncertainty of dose mapping caused by uncertainties in b-spline registration. Med Phys 2012;39:2186–92. doi:10.1118/1.3697524.

[231] Risholm P, Balter J, Wells WM. Estimation of delivered dose in radiotherapy: The influence of registration uncertainty. Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2011;6891 LNCS:548–55. doi:10.1007/978-3-642-23623-5_69.

[232] Tilly D, Tilly N, Ahnesjö A. Dose mapping sensitivity to deformable registration uncertainties in fractionated radiotherapy - applied to prostate proton treatments. BMC Med Phys 2013;13. doi:10.1186/1756-6649-13-2.

[233] Wang Y, Petit SF, Osorio EV, Gupta V, Romero AM, Heijmen B. An individualized strategy to estimate the effect of deformable registration uncertainty on accumulated dose in the upper abdomen. Phys Med Biol 2018;63. doi:10.1088/1361-6560/aac5c2.

[234] Rosu M, Chetty IJ, Balter JM, Kessler ML, McShan DL, Ten Haken RK. Dose reconstruction in deforming lung anatomy: Dose grid size effects and clinical implications. Med Phys 2005;32:2487–95. doi:10.1118/1.1949749.

[235] Yan C, Hugo G, Salguero FJ, Saleh-Sayah N, Weiss E, Sleeman WC, et al. A method to evaluate dose errors introduced by dose mapping processes for mass conserving deformations. Med Phys 2012;39:2119–28. doi:10.1118/1.3684951.

[236] Chetty IJ, Fontenot J. Adaptive Radiation Therapy: Off-Line, On-Line, and In-Line? Int J Radiat Oncol Biol Phys 2017;99:689–91. doi:10.1016/j.ijrobp.2017.07.017.

[237] Sonke JJ, Belderbos J. Adaptive Radiotherapy for Lung Cancer. Semin Radiat Oncol 2010;20:94–106. doi:10.1016/j.semradonc.2009.11.003.

[238] Marur S, Forastiere AA. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. Mayo Clin Proc 2016;91:386–96. doi:10.1016/j.mayocp.2015.12.017.

[239] Mermod M, Tolstonog G, Simon C, Monnier Y. Extracapsular spread in head and neck squamous cell carcinoma: A systematic review and meta-analysis. Oral Oncol 2016;62:60–71. doi:10.1016/j.oraloncology.2016.10.003.

[240] De Felice F, Thomas C, Barrington S, Pathmanathan A, Lei M, Urbano TG. Analysis of loco-regional failures in head and neck cancer after radical radiation therapy. Oral Oncol 2015;51:1051–5. doi:10.1016/j.oraloncology.2015.08.004.

[241] Zukauskaite R, Hansen CR, Brink C, Johansen J, Asmussen JT, Grau C, et al. Analysis of CT-verified loco-regional recurrences after definitive IMRT for HNSCC using site of origin estimation methods. Acta Oncol (Madr) 2017;56:1554–61. doi:10.1080/0284186X.2017.1346384.

[242] McCulloch MM, Muenz DG, Schipper MJ, Velec M, Dawson LA, Brock KK. A simulation study to assess the potential impact of developing normal tissue complication probability models with accumulated dose. Adv Radiat Oncol 2018;3:662–72. doi:10.1016/j.adro.2018.05.003.

[243] Berkovic P, Paelinck L, Lievens Y, Gulyban A, Goddeeris B, Derie C, et al. Adaptive radiotherapy for locally advanced non-small cell lung cancer, can we predict when and for whom? Acta Oncol (Madr) 2015;54:1438–44. doi:10.3109/0284186X.2015.1061209.

[244] Green OL, Henke LE, Hugo GD. Practical Clinical Workflows for Online and Offline Adaptive Radiation Therapy. Semin Radiat Oncol 2019;29:219–27. doi:10.1016/j.semradonc.2019.02.004.

[245] Lim-Reinders S, Keller BM, Al-Ward S, Sahgal A, Kim A. Online Adaptive Radiation Therapy. Int J Radiat Oncol Biol Phys 2017;99:994–1003. doi:10.1016/j.ijrobp.2017.04.023.

[246] Qin A, Sun Y, Liang J, Yan D. Evaluation of online/offline image guidance/adaptation approaches for prostate cancer radiation therapy. Int J Radiat Oncol Biol Phys 2015;91:1026–33. doi:10.1016/j.ijrobp.2014.12.043.

[247] Cheo T, Loh Y, Chen D, Lee KM, Tham I. Measuring radiotherapy setup errors at multiple neck levels in nasopharyngeal cancer (NPC): A case for differential PTV expansion. Radiother Oncol 2015;117:419–24. doi:10.1016/j.radonc.2015.09.032.

[248] Marks LB, Adams RD, Pawlicki T, Blumberg AL, Hoopes D, Brundage MD, et al. Enhancing the role of case-oriented peer review to improve quality and safety in radiation oncology: Executive summary. Pract Radiat Oncol 2013;3:149–56. doi:10.1016/j.prro.2012.11.010.

[249] Lo AC, Liu M, Chan E, Lund C, Truong PT, Loewen S, et al. The impact of peer review of volume delineation in stereotactic body radiation therapy planning for primary lung cancer: A multicenter quality assurance study. J Thorac Oncol 2014;9:527–33. doi:10.1097/JTO.0000000000000119.

[250] Verduijn GM, Bartels LW, Raaijmakers CPJ, Terhaard CHJ, Pameijer FA, van den Berg CAT. Magnetic Resonance Imaging Protocol Optimization for Delineation of Gross Tumor Volume in Hypopharyngeal and Laryngeal Tumors. Int J Radiat Oncol Biol Phys 2009;74:630–6. doi:10.1016/j.ijrobp.2009.01.014.

[251] Schinagl DAX, Span PN, Van Den Hoogen FJA, Merkx MAW, Slootweg PJ, Oyen WJG, et al. Pathology-based validation of FDG PET segmentation tools for volume assessment of lymph node metastases from head and neck cancer. Eur J Nucl Med Mol Imaging 2013;40:1828–35. doi:10.1007/s00259-013-2513-9.

[252] Bird D, Scarsbrook AF, Sykes J, Ramasamy S, Subesinghe M, Carey B, et al. Multimodality imaging with CT, MR and FDG-PET for radiotherapy target volume delineation in oropharyngeal squamous cell carcinoma. BMC Cancer 2015;15:1–10. doi:10.1186/s12885-015-1867-8.

[253] Schakel T, Hoogduin JM, Terhaard CHJ, Philippens MEP. Diffusion weighted MRI in head-and-neck cancer: Geometrical accuracy. Radiother Oncol 2013;109:394–7. doi:10.1016/j.radonc.2013.10.004.

[254] Paulson ES, Erickson B, Schultz C, Allen Li X. Comprehensive MRI simulation methodology using a dedicated MRI scanner in radiation oncology for external beam radiation treatment planning. Med Phys 2015;42:28–39. doi:10.1118/1.4896096.

[255] Webster GJ, Kilgallon JE, Ho KF, Rowbottom CG, Slevin NJ, Mackay RI. A novel imaging technique for fusion of high-quality immobilised MR images of the head and neck with CT scans for radiotherapy target delineation. Br J Radiol 2009;82:497–503. doi:10.1259/bjr/50709041.

[256] Hanvey S, McJury M, Tho LM, Glegg M, Thomson M, Grose D, et al. The influence of MRI scan position on patients with oropharyngeal cancer undergoing radical radiotherapy. Radiat Oncol 2013;8:1–10. doi:10.1186/1748-717X-8-129.

[257] Chuter R, Prestwich R, Bird D, Scarsbrook A, Sykes J, Wilson D, et al. The use of deformable image registration to integrate diagnostic MRI into the radiotherapy planning pathway for head and neck cancer. Radiother Oncol 2017;122:229–35. doi:10.1016/j.radonc.2016.07.016.

[258] Huang B, Chen Z, Wu PM, Ye Y, Feng ST, Wong CYO, et al. Fully Automated Delineation of Gross Tumor Volume for Head and Neck Cancer on PET-CT Using Deep Learning: A Dual-Center Study. Contrast Media Mol Imaging 2018;2018. doi:10.1155/2018/8923028.

[259] Thompson RF, Valdes G, Fuller CD, Carpenter CM, Morin O, Aneja S, et al. Artificial intelligence in radiation oncology: A specialty-wide disruptive transformation? Radiother Oncol 2018;129:421–6. doi:10.1016/j.radonc.2018.05.030.

[260] Arimura H, Soufi M, Kamezawa H, Ninomiya K, Yamada M. Radiomics with artificial intelligence for precision medicine in radiation therapy. J Radiat Res 2019;60:150–7. doi:10.1093/jrr/rry077.

[261] Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks: Med Phys 2017;44:547–57. doi:10.1002/mp.12045.

[262] Delpon G, Escande A, Ruef T, Darréon J, Fontaine J, Noblet C, et al. Comparison of automated atlas-based segmentation software for postoperative prostate cancer radiotherapy. Front Oncol 2016;6:1–6. doi:10.3389/fonc.2016.00178.

[263] Lustberg T, van Soest J, Gooding M, Peressutti D, Aljabar P, van der Stoep J, et al. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. Radiother Oncol 2018;126:312–7. doi:10.1016/j.radonc.2017.11.012.

[264] Jarrett D, Stride E, Vallis K, Gooding MJ. Applications and limitations of machine learning in radiation oncology. Br J Radiol 2019;92:1–12. doi:10.1259/bjr.20190001.

[265] Dawson LA, Anzai Y, Marsh L, Martel MK, Paulino A, Ship JA, et al. Patterns of local-regional recurrence following parotid-sparing conformal and segmental intensity-modulated radiotherapy for head and neck cancer. Int J Radiat Oncol Biol Phys 2000;46:1117–26. doi:10.1016/S0360-3016(99)00550-7.

[266] Chen AM, Farwell DG, Luu Q, Chen LM, Vijayakumar S, Purdy JA. Marginal misses after postoperative intensity-modulated radiotherapy for head and neck cancer. Int J Radiat Oncol Biol Phys 2011;80:1423–9. doi:10.1016/j.ijrobp.2010.04.011.

[267] Fiedler M, Weber F, Hautmann MG, Haubner F, Reichert TE, Klingelhöffer C, et al. Biological predictors of radiosensitivity in head and neck squamous cell carcinoma. Clin Oral Investig 2018;22:189–200. doi:10.1007/s00784-017-2099-x.

[268] Akervall J, Nandalur S, Zhang J, Qian CN, Goldstein N, Gyllerup P, et al. A novel panel of biomarkers predicts radioresistance in patients with squamous cell carcinoma of the head and neck. Eur J Cancer 2014;50:570–81. doi:10.1016/j.ejca.2013.11.007.

[269] Kanehira T, van Kranen S, Sonke JJ. PO-0993 Uncertainty estimation of dose accumulation with deformable image registration in head and neck region. Radiother. Oncol. 133, 2019, p. S546–7.

[270] Raaymakers BW, Jürgenliemk-Schulz IM, Bol GH, Glitzner M, Kotte ANTJ, Van Asselen B, et al. First patients treated with a 1.5 T MRI-Linac: Clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment. Phys Med Biol 2017;62:L41–50. doi:10.1088/1361-6560/aa9517.

[271] Bol GH, Hissoiny S, Lagendijk JJW, Raaymakers BW. Fast online Monte Carlo-based IMRT planning for the MRI linear accelerator. Phys Med Biol 2012;57:1375–85. doi:10.1088/0031-9155/57/5/1375.

[272] Deutschmann H, Kametriser G, Steininger P, Scherer P, Schöller H, Gaisberger C, et al. First clinical release of an online, adaptive, aperture-based image-guided radiotherapy strategy in intensity-modulated radiotherapy to correct for inter- and intrafractional rotations of the prostate. Int J Radiat Oncol Biol Phys 2012;83:1624–32. doi:10.1016/j.ijrobp.2011.10.009.

[273] Fischer-Valuck BW, Henke L, Green O, Kashani R, Acharya S, Bradley JD, et al. Two-and-a-half-year clinical experience with the world's first magnetic resonance image guided radiation therapy system. Adv Radiat Oncol 2017;2:485–93. doi:10.1016/j.adro.2017.05.006.

[274] Teske H, Bartelheimer K, Meis J, Bendl R, Stoiber EM, Giske K. Construction of a biomechanical head and neck motion model as a guide to evaluation of deformable image registration. Phys Med Biol 2017;62:N271–84. doi:10.1088/1361-6560/aa69b6.

[275] Al-Mayah A, Moseley J, Velec M, Brock K. Toward efficient biomechanical-based deformable image registration of lungs for image-guided radiotherapy. Phys Med Biol 2011;56:4701–13. doi:10.1088/0031-9155/56/15/005.

[276] Brown E, Owen R, Harden F, Mengersen K, Oestreich K, Houghton W, et al. Predicting the need for adaptive radiotherapy in head and neck cancer. Radiother Oncol 2015;116:57–63. doi:10.1016/j.radonc.2015.06.025.

[277] Murdoch-Kinch CA, Kim HM, Vineberg KA, Ship JA, Eisbruch A. Dose-Effect Relationships for the Submandibular Salivary Glands and Implications for Their Sparing by Intensity Modulated Radiotherapy. Int J Radiat Oncol Biol Phys 2008;72:373–82. doi:10.1016/j.ijrobp.2007.12.033.

APPENDIX A: SUPPLEMENTARY DATA

A.1: Patient group

Twelve patients with cancers in the head-and-neck region that were previously treated radically were selected for this research. The planning computed tomography (*pCT*) scans (Brilliance Big Bore; Philips Medical Systems, Eindhoven, The Netherlands) had a voxel size of 1.3 x 1.3 x 3.0 mm^3. Patients were immobilized using a 2.4 mm Reloadable Head and Shoulder S-Frame Kevlar Mask (Q-Fix, Avondale PA, USA) and an individual head and shoulder support vacuum bag (Klarity Medical Products, Newark OH, USA). Daily (n = 30) cone beam computed tomography (CBCT) scans were acquired to verify patient treatment position. The CBCT scans with a voxel size of 0.5 x 0.5 x 2 mm^3 were acquired prior to treatment using a Varian Truebeam (v2.0 or v2.5; Varian Medical Systems, Palo Alto CA, USA). For a minority of treatment fractions, CBCT scans were acquired on a Varian Clinac (v2.1) with a voxel size of 0.7 x 0.7 x 2.5 mm^3. All patients consented to their data being used for retrospective audits.

Table A.1: Clinical and disease characteristics of patient cohort.

Characteristic	Number
Number of patients	12
Primary site	
Tonsil	7
Nasopharynx	3
Oropharynx	2
Glomus Jugulare (treated to 54 Gy)	1
T stage	
T1	2
T2	5
T3	2
T4	2
No TNM	1
Concurrent chemotherapy	
Yes	7
No	5
Weight at treatment planning *(kg)*	
Range	53.2 – 141.3
Mean	75.2
Age *(years)*	
Range	40 – 80
Mean	62.8
Sex	
Male	7
Female	5

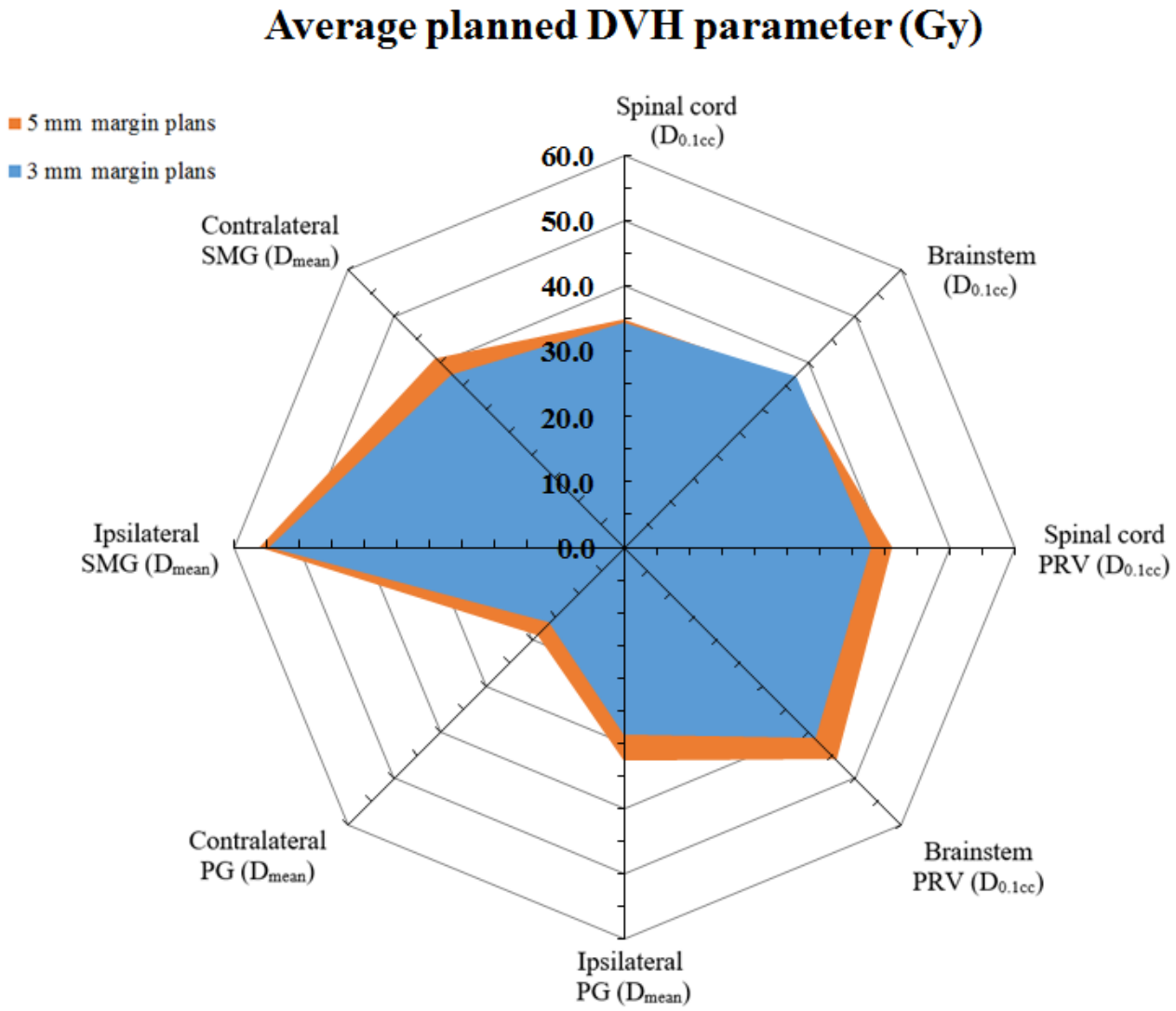


Figure A.1: Average planned dose-volume histogram (DVH) parameter over all patients with either 3 or 5 mm planning target volume (PTV) and planning risk volume (PRV) margins. PG = parotid gland. SMG = submandibular gland. $D_{0.1cc}$ = maximum dose to 0.1 cubic centimeters. D_{mean}= mean dose

Table A.2: Patient-specific dose volume histogram (DVH) metrics detailing the planned dose in Gy for each structure of interest using 3 mm planning target volume (PTV) margins. CTV = clinical target volume. PRV = planning risk volume. PG = parotid gland. SMG = submandibular gland. $D_{99\%}$ = minimum dose to 99% of volume. $D_{98\%}$ = minimum dose to 98% of volume. D_{mean} = mean dose. $D_{0.1cc}$ = maximum dose to 0.1 cubic centimeters.

	Patient	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***median***	***range***
Structure	***DVH parameter***														
CTV 54	$D_{99\%}$	53.5	52.7	53.4	53.9	52.3	54.1	53.2	53.5	54.1	53.8	53.6	53.0	**53.5**	52.3 – 54.1
CTV 60	$D_{99\%}$						60.0			60.2	60.4		59.0	**60.1**	59.0 – 60.4
CTV 66	$D_{99\%}$	64.8		64.7	65.2	65.1	65.7	64.6	65.4	65.5	64.4	65.2	63.6	**65.1**	63.6 – 65.7
PTV 54	$D_{98\%}$	52.2	51.5	51.5	51.3	51.8	53.0	51.7	53.4	51.5	51.3	52.2	51.7	**51.7**	51.3 – 53.4
PTV 60	$D_{98\%}$						58.2			57.7	57.1		57.4	**57.6**	57.1 – 58.2
PTV 66	$D_{98\%}$	63.8		63.6	63.1	64.2	63.5	62.9	63.7	64.1	62.7	63.3	63.2	**63.5**	62.7 – 64.2
Spinal cord	$D_{0.1cc}$	43.4	19.4	38.8	31.4	34.2	40.5	38.9	34.1	33.7	31.4	36.0	32.3	**34.1**	19.4 – 43.4
Brainstem	$D_{0.1cc}$	35.7	51.9	37.9	30.6	38.0	30.0	39.1	35.8	38.5	37.1	37.1	35.2	**37.1**	30.0 – 51.9
Spinal cord PRV	$D_{0.1cc}$	44.5	21.1	40.5	34.8	34.8	45.2	41.4	45.2	39.0	34.5	38.4	35.3	**38.7**	21.1 – 45.2
Brainstem PRV	$D_{0.1cc}$	38.3	52.9	43.6	33.7	40.3	36.0	45.3	36.9	41.6	45.7	41.4	41.5	**41.5**	33.7 – 52.9
Ipsilateral PG	D_{mean}	27.4	14.1	24.4	24.6	34.4	30.8	36.2	34.5	28.6	20.2	27.2	43.6	**28.0**	14.1 – 43.6
								36.2		20.1					
Contralateral PG	D_{mean}	25.2	4.5	25.9	24.0	32.8	6.5		4.8		18.0	16.1	4.6	**17.1**	4.5 – 32.8
Ipsilateral SMG	D_{mean}	64.7	1.1	58.4	63.5	63.7	63.5	42.2	64.6	65.8	52.4		64.5	**63.5**	1.1 – 65.8
								43.6		58.5					
Contralateral SMG	D_{mean}	51.6	0.7	55.4	50.7	55.5	24.0		27.6		44.3	51.7		**50.7**	0.7 – 55.5

Table A.3: Patient-specific dose volume histogram (DVH) metrics detailing the planned dose in Gy for each structure of interest using 5 mm planning target volume (PTV) margins. CTV = clinical target volume. PRV = planning risk volume. PG = parotid gland. SMG = submandibular gland. $D_{99\%}$ = minimum dose to 99% of volume. $D_{98\%}$ = minimum dose to 98% of volume. D_{mean} = mean dose. $D_{0.1cc}$ = maximum dose to 0.1 cubic centimeters.

	Patient	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***median***	***range***
Structure	***DVH parameter***														
CTV 54	$D_{99\%}$	53.4	52.5	53.6	54.0	52.8	53.6	54.3	53.8	54.6	54.0	54.0	53.2	**53.7**	52.5 – 54.6
CTV 60	$D_{99\%}$						59.6			60.0	60.0		60.2	**60.0**	59.6 – 60.2
CTV 66	$D_{99\%}$	64.3		65.0	65.2	65.1	65.1	64.6	65.8	64.6	64.2	65.9	63.7	**65.0**	63.7 – 65.9
PTV 54	$D_{98\%}$	52.0	51.5	51.6	52.7	51.4	53.0	51.9	53.5	52.4	52.0	52.6	51.6	**52.0**	51.4 – 53.5
PTV 60	$D_{98\%}$						57.7			57.6	57.2		57.5	**57.6**	57.2 – 57.7
PTV 66	$D_{98\%}$	63.4		63.5	63.6	63.6	62.9	62.8	63.1	63.2	62.8	63.6	63.6	**63.4**	62.8 – 63.6
Spinal cord	$D_{0.1cc}$	42.3	23.5	37.5	32.7	32.2	36.5	42.0	35.6	29.1	34.0	36.5	37.6	**36.0**	23.5 – 42.3
Brainstem	$D_{0.1cc}$	42.1	52.4	34.2	30.6	36.6	27.7	27.1	36.5	37.9	45.8	34.7	32.6	**35.6**	27.1 – 52.4
Spinal cord PRV	$D_{0.1cc}$	45.3	26.6	42.0	37.8	36.9	42.3	47.4	42.3	40.8	44.4	41.4	46.5	**42.1**	26.6 – 47.4
Brainstem PRV	$D_{0.1cc}$	44.3	53.3	47.3	36.5	41.2	42.6	46.2	42.6	46.3	60.7	43.7	46.5	**45.3**	36.5 – 60.7
Ipsilateral PG	D_{mean}	30.8	16.8	27.1	27.2	36.2	33.0	42.3 46.2	35.3	26.3 25.5	26.1	35.9	50.0	**31.9**	16.8 – 50.0
Contralateral PG	D_{mean}	26.8	6.4	26.7	26.1	37.2	8.9		5.2		22.1	21.5	8.6	**21.8**	5.2 – 37.2
Ipsilateral SMG	D_{mean}	64.1	1.5	59.2	65.2	64.4	64.4	47.4 48.0	66.1	65.2 61.5	56.3		66.0	**64.1**	1.5 – 66.1
Contralateral SMG	D_{mean}	51.3	0.9	55.7	53.4	56.0	26.5		27.6		49.3	54.4		**51.3**	0.9 – 56.0

A.2: Uncertainty analysis

Supplementary Data A.2 details sources of uncertainty that were identified in this work and how they were assessed. For complete comprehension of the uncertainty quantification, the reader is directed to the GUM guidelines [196].

List of abbreviations

BA = bony anatomy (e.g., C1-C3, C3-C5, C5-C7, mandible, etc…)

PG = parotid gland

DIR = deformable image registration

MMR = manual method of registration

pCT = planning computed tomography

dCT = deformed computed tomography

CBCT = cone beam computed tomography

HR = high resolution

SD = standard deviation

N = sample size

Parameters used in this work following GUM guidelines

D_f^{BA} = the measurand estimating *BA/PG* deviation relative to C1-C3 at fraction *f*

V_f^{PG} = the measurand estimating *PG/neck* volume at fraction *f*

$c(T_{x,f}^{BA})$ = sensitivity coefficient for change in centroid position of *BA/PG* in *x* plane at fraction *f*

$c(V^{PG})$ = sensitivity coefficient for volume of *PG/neck* structure

$u(T_{x,a}^{BA})$ = uncertainty component *a* for change in centroid position of *BA/PG* in *x* plane

$u(V_a^{PG})$ = uncertainty component *a* for volume of *PG/neck* structure

$u_c(T_x^{BA})$ = **combined** uncertainty for change in centroid position of *BA/PG* in *x* plane

$u_c(V^{PG})$ = **combined** uncertainty for volume of *PG/neck* structure

$u_{o,f}(D_f^{BA})$ = **overall** uncertainty for *BA/PG* deviation measurand at fraction *f*

$u_{o,f}(V_f^{PG})$ = **overall** uncertainty for *PG/neck* volume measurand at fraction *f*

k_p = coverage factor determining level of confidence

List of formula

$$D_f^{BA} = \sqrt{\left(T_{x,f}^{BA} - T_{x,f}^{C1-C3}\right)^2 + \left(T_{y,f}^{BA} - T_{y,f}^{C1-C3}\right)^2 + \left(T_{z,f}^{BA} - T_{z,f}^{C1-C3}\right)^2}$$

V_f^{PG} = volume recorded in MATLAB analysis

$$c(T_{x,f}^{BA}) \equiv \partial D_f^{BA}/\partial T_{x,f}^{BA} = \partial D_f^{BA}/\partial T_{x,f}^{C1-C3} = \frac{\left(T_{x,f}^{BA} - T_{x,f}^{C1-C3}\right)}{\sqrt{\left(T_{x,f}^{BA} - T_{x,f}^{C1-C3}\right)^2 + \left(T_{y,f}^{BA} - T_{y,f}^{C1-C3}\right)^2 + \left(T_{z,f}^{BA} - T_{z,f}^{C1-C3}\right)^2}}$$

$$c(T_{y,f}^{BA}) \equiv \partial D_f^{BA}/\partial T_{y,f}^{BA} = \partial D_f^{BA}/\partial T_{y,f}^{C1-C3} = \frac{\left(T_{y,f}^{BA} - T_{y,f}^{C1-C3}\right)}{\sqrt{\left(T_{x,f}^{BA} - T_{x,f}^{C1-C3}\right)^2 + \left(T_{y,f}^{BA} - T_{y,f}^{C1-C3}\right)^2 + \left(T_{z,f}^{BA} - T_{z,f}^{C1-C3}\right)^2}}$$

$$c(T_{z,f}^{BA}) \equiv \partial D_f^{BA}/\partial T_{z,f}^{BA} = \partial D_f^{BA}/\partial T_{z,f}^{C1-C3} = \frac{\left(T_{z,f}^{BA} - T_{z,f}^{C1-C3}\right)}{\sqrt{\left(T_{x,f}^{BA} - T_{x,f}^{C1-C3}\right)^2 + \left(T_{y,f}^{BA} - T_{y,f}^{C1-C3}\right)^2 + \left(T_{z,f}^{BA} - T_{z,f}^{C1-C3}\right)^2}}$$

$$c(V^{PG}) \equiv \partial V^{PG}/\partial V^{PG} = 1$$

The **combined** uncertainty was calculated for each measurand input variable (i.e., $T_{x,f}^{BA}, T_{z,f}^{C1-C3}, etc$ ...) and are detailed in the following tables of Supplementary Data B. The **overall** uncertainty for each fraction *f*'s measurand was then calculated using:

$$u_{o,f}\left(D_f^{BA}\right) = \sum_{i=1}^{N} c\left(T_{i,f}\right) . u_c(T_i)$$

The **overall** uncertainty for each fraction was then multiplied by a coverage factor of 1.645 (90% level of confidence) to give an **expanded** uncertainty. The coverage factor was chosen after Shapiro–Wilk tests were unable to reject normality of the measurands' distributions. Finally, the **expanded** uncertainty was used as the uncertainty associated with the metrics of interest used in Equation 3 (page 79) of the dissertation. Note that **combined** uncertainties of each measurand input variable are identical for every fraction. The fraction-varying sensitivity coefficients give rise to different **overall** uncertainties for the same BA region of the same patient.

The uncertainties associated with μ_0 and σ of Equation 4 (page 79) of the dissertation were the standard error of the reference period and the standard error of the reference period's SD multiplied by a coverage factor of 1.645 (90% level of confidence), respectively.

Uncertainties in metrics of interest derived with DIR and MMR are established following the same methodology as described above, however note the differences in uncertainty components in the following tables.

List of tables

Tables A.2.1.X

Table A.2.1.1: DIR uncertainty components for change in centroid position of C1-C3 in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{C1-C3})$	Uncertainty in DIR algorithm	0.18	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{C1-C3})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.16	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{C1-C3})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.18	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{C1-C3})$	Conversion to HR for dCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{C1-C3})$	Conversion to HR for pCT structure	0.04	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{C1-C3})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{C1-C3})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{C1-C3})$	Uncertainty in DIR algorithm	0.11	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{C1-C3})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.10	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{C1-C3})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.12	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{C1-C3})$	Conversion to HR for dCT structure	0.03	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{C1-C3})$	Conversion to HR for pCT structure	0.03	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{C1-C3})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{C1-C3})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{C1-C3})$	Uncertainty in DIR algorithm	0.25	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{C1-C3})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.31	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{C1-C3})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.32	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{C1-C3})$	Conversion to HR for dCT structure	0.06	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{C1-C3})$	Conversion to HR for pCT structure	0.04	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{C1-C3})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{C1-C3})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{C1-C3}) = 0.31\ mm}$ $\boldsymbol{u_c(T_y^{C1-C3}) = 0.20\ mm}$ $\boldsymbol{u_c(T_z^{C1-C3}) = 0.52\ mm}$

Table A.2.1.2: DIR uncertainty components for change in centroid position of C3-C5 in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{C3-C5})$	Uncertainty in DIR algorithm	0.11	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{C3-C5})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.20	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{C3-C5})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.23	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{C3-C5})$	Conversion to HR for dCT structure	0.06	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{C3-C5})$	Conversion to HR for pCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{C3-C5})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{C3-C5})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{C3-C5})$	Uncertainty in DIR algorithm	0.35	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{C3-C5})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.09	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{C3-C5})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.19	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{C3-C5})$	Conversion to HR for dCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{C3-C5})$	Conversion to HR for pCT structure	0.03	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{C3-C5})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{C3-C5})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{C3-C5})$	Uncertainty in DIR algorithm	0.23	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{C3-C5})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.19	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{C3-C5})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.34	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{C3-C5})$	Conversion to HR for dCT structure	0.07	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{C3-C5})$	Conversion to HR for pCT structure	0.03	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{C3-C5})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{C3-C5})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{C3-C5}) = 0.33\ mm}$ $\boldsymbol{u_c(T_y^{C3-C5}) = 0.41\ mm}$ $\boldsymbol{u_c(T_z^{C3-C5}) = 0.46\ mm}$

Table A.2.1.3: DIR uncertainty components for change in centroid position of C5-C7 in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{C5-C7})$	Uncertainty in DIR algorithm	0.08	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{C5-C7})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.28	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{C5-C7})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.31	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{C5-C7})$	Conversion to HR for dCT structure	0.06	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{C5-C7})$	Conversion to HR for pCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{C5-C7})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{C5-C7})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{C5-C7})$	Uncertainty in DIR algorithm	0.44	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{C5-C7})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.21	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{C5-C7})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.28	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{C5-C7})$	Conversion to HR for dCT structure	0.04	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{C5-C7})$	Conversion to HR for pCT structure	0.04	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{C5-C7})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{C5-C7})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{C5-C7})$	Uncertainty in DIR algorithm	0.41	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{C5-C7})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.33	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{C5-C7})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.30	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{C5-C7})$	Conversion to HR for dCT structure	0.04	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{C5-C7})$	Conversion to HR for pCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{C5-C7})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{C5-C7})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{C5-C7}) = 0.44\ mm}$ $\boldsymbol{u_c(T_y^{C5-C7}) = 0.57\ mm}$ $\boldsymbol{u_c(T_z^{C5-C7}) = 0.61\ mm}$

Table A.2.1.4: DIR uncertainty components for change in centroid position of mandible in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{mandible})$	Uncertainty in DIR algorithm	0.62	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{mandible})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.10	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{mandible})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.28	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{mandible})$	Conversion to HR for dCT structure	0.11	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{mandible})$	Conversion to HR for pCT structure	0.06	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{mandible})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{mandible})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{mandible})$	Uncertainty in DIR algorithm	0.47	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{mandible})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.34	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{mandible})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.26	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{mandible})$	Conversion to HR for dCT structure	0.09	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{mandible})$	Conversion to HR for pCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{mandible})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{mandible})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{mandible})$	Uncertainty in DIR algorithm	0.40	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{mandible})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.21	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{mandible})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.25	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{mandible})$	Conversion to HR for dCT structure	0.07	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{mandible})$	Conversion to HR for pCT structure	0.03	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{mandible})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{mandible})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{mandible}) = 0.70\ mm}$ $\boldsymbol{u_c(T_y^{mandible}) = 0.64\ mm}$ $\boldsymbol{u_c(T_z^{mandible}) = 0.52\ mm}$

Table A.2.1.5: DIR uncertainty components for change in centroid position of maxilla in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{maxilla})$	Uncertainty in DIR algorithm	0.43	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{maxilla})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.34	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{maxilla})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.56	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{maxilla})$	Conversion to HR for dCT structure	0.13	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{maxilla})$	Conversion to HR for pCT structure	0.14	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{maxilla})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{maxilla})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{maxilla})$	Uncertainty in DIR algorithm	0.55	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{maxilla})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.36	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{maxilla})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.30	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{maxilla})$	Conversion to HR for dCT structure	0.07	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{maxilla})$	Conversion to HR for pCT structure	0.06	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{maxilla})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{maxilla})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{maxilla})$	Uncertainty in DIR algorithm	1.02	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{maxilla})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.70	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{maxilla})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.34	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{maxilla})$	Conversion to HR for dCT structure	0.08	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{maxilla})$	Conversion to HR for pCT structure	0.08	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{maxilla})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{maxilla})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{maxilla}) = 0.80\ mm}$ $\boldsymbol{u_c(T_y^{maxilla}) = 0.72\ mm}$ $\boldsymbol{u_c(T_z^{maxilla}) = 1.29\ mm}$

Table A.2.1.6: DIR uncertainty components for change in centroid position of base of skull (BOS) in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{BOS})$	Uncertainty in DIR algorithm	0.47	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{BOS})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.36	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{BOS})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.70	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{BOS})$	Conversion to HR for dCT structure	0.13	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{BOS})$	Conversion to HR for pCT structure	0.12	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{BOS})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{BOS})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{BOS})$	Uncertainty in DIR algorithm	0.44	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{BOS})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.86	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{BOS})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.34	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{BOS})$	Conversion to HR for dCT structure	0.06	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{BOS})$	Conversion to HR for pCT structure	0.06	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{BOS})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{BOS})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{BOS})$	Uncertainty in DIR algorithm	0.90	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{BOS})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.34	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{BOS})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.30	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{BOS})$	Conversion to HR for dCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{BOS})$	Conversion to HR for pCT structure	0.03	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{BOS})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{BOS})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{BOS}) = 0.93\ mm}$ $\boldsymbol{u_c(T_y^{BOS}) = 1.03\ mm}$ $\boldsymbol{u_c(T_z^{BOS}) = 1.01\ mm}$

Table A.2.1.7: DIR uncertainty components for change in centroid position of hyoid in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{hyoid})$	Uncertainty in DIR algorithm	0.90	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{hyoid})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.42	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{hyoid})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	1.07	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{hyoid})$	Conversion to HR for dCT structure	0.22	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{hyoid})$	Conversion to HR for pCT structure	0.21	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{hyoid})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{hyoid})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{hyoid})$	Uncertainty in DIR algorithm	0.77	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{hyoid})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.60	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{hyoid})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.53	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{hyoid})$	Conversion to HR for dCT structure	0.13	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{hyoid})$	Conversion to HR for pCT structure	0.13	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{hyoid})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{hyoid})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{hyoid})$	Uncertainty in DIR algorithm	0.10	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{hyoid})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.37	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{hyoid})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.19	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{hyoid})$	Conversion to HR for dCT structure	0.04	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{hyoid})$	Conversion to HR for pCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{hyoid})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{hyoid})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{hyoid}) = 1.49\ mm}$ $\boldsymbol{u_c(T_y^{hyoid}) = 1.13\ mm}$ $\boldsymbol{u_c(T_z^{hyoid}) = 0.43\ mm}$

Table A.2.1.8: DIR uncertainty components for change in centroid position of occipital in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{occipital})$	Uncertainty in DIR algorithm	1.72	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{occipital})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.21	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{occipital})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.51	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{occipital})$	Conversion to HR for dCT structure	0.15	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{occipital})$	Conversion to HR for pCT structure	0.14	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{occipital})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{occipital})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{occipital})$	Uncertainty in DIR algorithm	1.00	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{occipital})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.24	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{occipital})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.39	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{occipital})$	Conversion to HR for dCT structure	0.11	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{occipital})$	Conversion to HR for pCT structure	0.06	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{occipital})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{occipital})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{occipital})$	Uncertainty in DIR algorithm	1.54	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{occipital})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.09	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{occipital})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.25	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{occipital})$	Conversion to HR for dCT structure	0.07	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{occipital})$	Conversion to HR for pCT structure	0.03	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{occipital})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{occipital})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{occipital}) = 1.82\ mm}$ $\boldsymbol{u_c(T_y^{occipital}) = 1.11\ mm}$ $\boldsymbol{u_c(T_z^{occipital}) = 1.57\ mm}$

Table A.2.1.9: DIR uncertainty components for change in centroid position of larynx in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{larynx})$	Uncertainty in DIR algorithm	0.38	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{larynx})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.37	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{larynx})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.60	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{larynx})$	Conversion to HR for dCT structure	0.29	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{larynx})$	Conversion to HR for pCT structure	0.21	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{larynx})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{larynx})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{larynx})$	Uncertainty in DIR algorithm	0.18	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{larynx})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.57	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{larynx})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.33	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{larynx})$	Conversion to HR for dCT structure	0.19	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{larynx})$	Conversion to HR for pCT structure	0.22	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{larynx})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{larynx})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{larynx})$	Uncertainty in DIR algorithm	0.08	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{larynx})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.23	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{larynx})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.23	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{larynx})$	Conversion to HR for dCT structure	0.09	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{larynx})$	Conversion to HR for pCT structure	0.11	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{larynx})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{larynx})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)

$\boldsymbol{u_c(T_x^{larynx}) = 0.88\ mm}$ $\boldsymbol{u_c(T_y^{larynx}) = 0.74\ mm}$ $\boldsymbol{u_c(T_z^{larynx}) = 0.37\ mm}$

Tables A.2.2.X

Table A.2.2.1: MMR uncertainty components for change in centroid position of C1-C3 in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{C1-C3})$	Inter-observer variation in match result	0.51	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{x,b}^{C1-C3})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{x,c}^{C1-C3})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,a}^{C1-C3})$	Inter-observer variation in match result	0.33	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{y,b}^{C1-C3})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,c}^{C1-C3})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,a}^{C1-C3})$	Inter-observer variation in match result	0.24	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{z,b}^{C1-C3})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,c}^{C1-C3})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u_c(T_x^{C1-C3}) = 0.64\ mm}$		$\boldsymbol{u_c(T_y^{C1-C3}) = 0.51\ mm}$	$\boldsymbol{u_c(T_z^{C1-C3}) = 0.90\ mm}$

Table A.2.2.2: MMR uncertainty components for change in centroid position of C3-C5 in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{C3-C5})$	Inter-observer variation in match result	0.19	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{x,b}^{C3-C5})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{x,c}^{C3-C5})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,a}^{C3-C5})$	Inter-observer variation in match result	0.34	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{y,b}^{C3-C5})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,c}^{C3-C5})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,a}^{C3-C5})$	Inter-observer variation in match result	0.20	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{z,b}^{C3-C5})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,c}^{C3-C5})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u_c(T_x^{C3-C5}) = 0.43\ mm}$		$\boldsymbol{u_c(T_y^{C3-C5}) = 0.51\ mm}$	$\boldsymbol{u_c(T_z^{C3-C5}) = 0.89\ mm}$

Table A.2.2.3: MMR uncertainty components for change in centroid position of C5-C7 in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{C5-C7})$	Inter-observer variation in match result	0.29	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{x,b}^{C5-C7})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{x,c}^{C5-C7})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,a}^{C5-C7})$	Inter-observer variation in match result	0.28	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{y,b}^{C5-C7})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,c}^{C5-C7})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,a}^{C5-C7})$	Inter-observer variation in match result	0.25	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{z,b}^{C5-C7})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,c}^{C5-C7})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u_c(T_x^{C5-C7}) = 0.48\ mm}$		$\boldsymbol{u_c(T_y^{C5-C7}) = 0.47\ mm}$	$\boldsymbol{u_c(T_z^{C5-C7}) = 0.90\ mm}$

Table A.2.2.4: MMR uncertainty components for change in centroid position of mandible in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{mandible})$	Inter-observer variation in match result	0.62	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{x,b}^{mandible})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{x,c}^{mandible})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,a}^{mandible})$	Inter-observer variation in match result	0.40	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{y,b}^{mandible})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,c}^{mandible})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,a}^{mandible})$	Inter-observer variation in match result	0.73	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{z,b}^{mandible})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,c}^{mandible})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u_c(T_x^{mandible}) = 0.73\ mm}$		$\boldsymbol{u_c(T_y^{mandible}) = 0.56\ mm}$	$\boldsymbol{u_c(T_z^{mandible}) = 1.13\ mm}$

Table A.2.2.5: MMR uncertainty components for change in centroid position of maxilla in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{maxilla})$	Inter-observer variation in match result	0.16	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{x,b}^{maxilla})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{x,c}^{maxilla})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,a}^{maxilla})$	Inter-observer variation in match result	0.15	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{y,b}^{maxilla})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,c}^{maxilla})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,a}^{maxilla})$	Inter-observer variation in match result	0.24	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{z,b}^{maxilla})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,c}^{maxilla})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u_c(T_x^{maxilla}) = 0.42\ mm}$		$\boldsymbol{u_c(T_y^{maxilla}) = 0.41\ mm}$	$\boldsymbol{u_c(T_z^{maxilla}) = 0.90\ mm}$

Table A.2.2.6: MMR uncertainty components for change in centroid position of base of skull (BOS) in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{BOS})$	Inter-observer variation in match result	0.12	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{x,b}^{BOS})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{x,c}^{BOS})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,a}^{BOS})$	Inter-observer variation in match result	0.12	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{y,b}^{BOS})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,c}^{BOS})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,a}^{BOS})$	Inter-observer variation in match result	0.25	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{z,b}^{BOS})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,c}^{BOS})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u_c(T_x^{BOS}) = 0.40\ mm}$		$\boldsymbol{u_c(T_y^{BOS}) = 0.40\ mm}$	$\boldsymbol{u_c(T_z^{BOS}) = 0.90\ mm}$

Table A.2.2.7: MMR uncertainty components for change in centroid position of hyoid in each anatomical plane.

Standard uncertainty component	*Source of uncertainty*	*Value (mm)*	*How uncertainty was assessed*
$u(T_{x,a}^{hyoid})$	Inter-observer variation in match result	0.25	SD of differences between match of two observers (authors NL and RL. N = 117)
$u(T_{x,b}^{hyoid})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$u(T_{x,c}^{hyoid})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$u(T_{y,a}^{hyoid})$	Inter-observer variation in match result	0.33	SD of differences between match of two observers (authors NL and RL. N = 117)
$u(T_{y,b}^{hyoid})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$u(T_{y,c}^{hyoid})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$u(T_{z,a}^{hyoid})$	Inter-observer variation in match result	0.41	SD of differences between match of two observers (authors NL and RL. N = 117)
$u(T_{z,b}^{hyoid})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$u(T_{z,c}^{hyoid})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$u_c(T_x^{hyoid}) = 0.46$ mm		$u_c(T_y^{hyoid}) = 0.51$ mm	$u_c(T_z^{hyoid}) = 0.96$ mm

Table A.2.2.8: MMR uncertainty components for change in centroid position of occipital in each anatomical plane.

Standard uncertainty component	*Source of uncertainty*	*Value (mm)*	*How uncertainty was assessed*
$u(T_{x,a}^{occipital})$	Inter-observer variation in match result	0.33	SD of differences between match of two observers (authors NL and RL. N = 117)
$u(T_{x,b}^{occipital})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$u(T_{x,c}^{occipital})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$u(T_{y,a}^{occipital})$	Inter-observer variation in match result	0.41	SD of differences between match of two observers (authors NL and RL. N = 117)
$u(T_{y,b}^{occipital})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$u(T_{y,c}^{occipital})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$u(T_{z,a}^{occipital})$	Inter-observer variation in match result	0.34	SD of differences between match of two observers (authors NL and RL. N = 117)
$u(T_{z,b}^{occipital})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$u(T_{z,c}^{occipital})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$u_c(T_x^{occipital}) = 0.50$ mm		$u_c(T_y^{occipital}) = 0.56$ mm	$u_c(T_z^{occipital}) = 0.93$ mm

Table A.2.2.9: MMR uncertainty components for change in centroid position of larynx in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{larynx})$	Inter-observer variation in match result	0.43	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{x,b}^{larynx})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{x grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{x,c}^{larynx})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,a}^{larynx})$	Inter-observer variation in match result	0.50	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{y,b}^{larynx})$	Intra-voxel location of centroid (function of grid resolution)	0.38	$\frac{\text{y grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{y,c}^{larynx})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,a}^{larynx})$	Inter-observer variation in match result	1.27	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(T_{z,b}^{larynx})$	Intra-voxel location of centroid (function of grid resolution)	0.87	$\frac{\text{z grid resolution}/2}{\sqrt{3}}$
$\boldsymbol{u}(T_{z,c}^{larynx})$	Rounding of treatment planning system value	< 0.01	$\frac{\text{rounding range}/2}{\sqrt{3}}$
$\boldsymbol{u_c(T_x^{larynx}) = 0.58\ mm}$		$\boldsymbol{u_c(T_y^{larynx}) = 0.63\ mm}$	$\boldsymbol{u_c(T_z^{larynx}) = 1.54\ mm}$

Tables A.2.3.X

Table A.2.3.1: DIR uncertainty components for change in centroid position of PGs in each anatomical plane.

Standard uncertainty component	***Source of uncertainty***	***Value (mm)***	***How uncertainty was assessed***
$\boldsymbol{u}(T_{x,a}^{PG})$	Uncertainty in DIR algorithm	0.11	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{x,b}^{PG})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.12	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,c}^{PG})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.12	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{x,d}^{PG})$	Conversion to HR for dCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,e}^{PG})$	Conversion to HR for pCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{x,f}^{PG})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{x,g}^{PG})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{x grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,a}^{PG})$	Uncertainty in DIR algorithm	0.10	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{y,b}^{PG})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.13	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,c}^{PG})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.29	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{y,d}^{PG})$	Conversion to HR for dCT structure	0.07	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,e}^{PG})$	Conversion to HR for pCT structure	0.05	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{y,f}^{PG})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{y,g}^{PG})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{y grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,a}^{PG})$	Uncertainty in DIR algorithm	0.10	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(T_{z,b}^{PG})$	Capping and other aspects (excl. grid size uncert.) when copying structure from pCT to CBCT	0.20	SD of differences between centroids of pCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,c}^{PG})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	0.27	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(T_{z,d}^{PG})$	Conversion to HR for dCT structure	0.07	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,e}^{PG})$	Conversion to HR for pCT structure	0.08	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u}(T_{z,f}^{PG})$	Grid size uncertainty of copying structure from pCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{pCT no.of points}}}$ (N = 120)
$\boldsymbol{u}(T_{z,g}^{PG})$	Grid size uncertainty of copying structure from dCT to CBCT	< 0.01	$\frac{(\text{z grid resolution of CBCT}/\sqrt{3})}{\sqrt{\text{dCT no.of points}}}$ (N = 120)
$\boldsymbol{u_c(T_x^{PG}) = 0.21\ mm}$		$\boldsymbol{u_c(T_y^{PG}) = 0.34\ mm}$	$\boldsymbol{u_c(T_z^{PG}) = 0.36\ mm}$

Tables A.2.4.X

Table A.2.4.1: DIR uncertainty components for change in volume of PGs

Standard uncertainty component	***Source of uncertainty***	***Value (cm^3)***	***How uncertainty was assessed***
$\boldsymbol{u}(V_a^{PG})$	Uncertainty in DIR algorithm	0.30	SD of differences between centroids of manually adapted DIR and DIR structures (N = 6)
$\boldsymbol{u}(V_b^{PG})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	1.37	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(V_c^{PG})$	Conversion to HR for dCT structure	0.04	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u_c(V^{PG}) = 1.40\ cm^3}$			

Table A.2.4.2: DIR uncertainty components for change in volume of neck

Standard uncertainty component	***Source of uncertainty***	***Value (cm^3)***	***How uncertainty was assessed***
$\boldsymbol{u}(V_a^{neck})$	Uncertainty in DIR algorithm	4.20	Estimated uncertainty of 2% of average neck volume over all patients
$\boldsymbol{u}(V_b^{neck})$	Capping and other aspects (excl. grid size uncert.) when copying structure from dCT to CBCT	8.09	SD of differences between centroids of dCT and CBCT structures correcting for grid size uncertainty (N = 30)
$\boldsymbol{u}(V_c^{neck})$	Conversion to HR for dCT structure	0.25	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u_c(V^{neck}) = 9.12\ cm^3}$			

Tables A.2.5.X

Table A.2.5.1: Manual uncertainty components for change in volume of a single axial slice at base of C2

Standard uncertainty component	***Source of uncertainty***	***Value (cm^3)***	***How uncertainty was assessed***
$\boldsymbol{u}(V_a^{C2})$	Inter-observer variation in volume definition	0.50	SD of differences between match of two observers (authors NL and RL. N = 117)
$\boldsymbol{u}(V_b^{C2})$	Conversion to HR for dCT structure	0.01	SD of differences between centroids of HR and non-HR structures (N = 24)
$\boldsymbol{u_c(V^{C2}) = 0.50\ cm^3}$			

A.3: SPC results

Table A.3.1: Fraction numbers when exponentially weighted moving average (EWMA) statistics exceeded process limits. EWMA statistics are calculated from data acquired using deformable image registration (DIR). (IC= in control; NA = not assessed; PG = parotid gland; BOS = base of skull; Occ = occipital bone; Max = maxilla; ↑/↓= increase/decrease in volume).

Patient number	D_f^{Occ}	D_f^{BOS}	D_f^{Max}	D_f^{C3-C5}	D_f^{C5-C7}	D_f^{Hyoid}	D_f^{Larynx}	$D_f^{Mandible}$	$D_f^{PG>26\ Gy}$	$D_f^{PG\leq 26\ Gy}$	$V_f^{PG>26\ Gy}$	$V_f^{PG\leq 26\ Gy}$	V_f^{neck}
1	13	30	IC	15	12	13	IC	9	11, 9	-	8, 9	-	IC
2	28	IC	9	NA	NA	NA	NA	15	-	9, 9	-	11, 10	10 ↑
3	IC	IC	IC	IC	30	16	IC	18	19, IC	-	10, 8	-	12 ↓
4	30	30	30	18	19	IC	18	22	30, IC	-	17, 15	-	17 ↑
5	13	20	11	15	NA	23	NA	11	14, 20	-	9, 9	-	10 ↓
6	19	20	NA	IC	IC	17	IC	17	14	22	8	9	24 ↓
7	25	IC	NA	IC	NA	23	25	IC	18	7	11	12	NA
8	IC	10	12	11	NA	IC	20	8	13, 15	-	20, 11	-	9 ↓
9	12	9	IC	IC	NA	11	20	IC	13	18	12	21	IC
10	IC	25	15	13	18	14	NA	17	18	24	9	20	14 ↓
11	IC	9	8	10	14	12	14	IC	IC, 11	-	21, 12	-	18 ↓
12	IC	20	23	14	13	16	IC	IC	12	9	8	12	15 ↓
IC%	**42%**	**25%**	**30%**	**36%**	**14%**	**18%**	**44%**	**33%**	**18%**	**0%**	**0%**	**0%**	**18%**
Median	**19**	**20**	**12**	**14**	**16**	**16**	**20**	**16**	**14**	**9**	**10**	**12**	**14**

A.4: Individual patient 90% level of confidence SPC results

Table A.4.1: Fraction number when exponentially weighted moving average (EWMA) statistics exceeded process limits before the end of treatment week 4 shown in boldface. Fraction number when the accuracy interval exceeded the process limits 90% level of confidence shown in parentheses. EWMA parameters are established from deformable image registration (DIR) raw data. (No entry = in control or exceeded control after treatment week 4; NA = not assessed; PG = parotid gland; BOS = base of skull; Occ = occipital; Max = maxilla; Mand = mandible; X = accuracy interval did not exceed process limits 90% level of confidence; ↑/↓= increase/decrease in volume).

Patient number	D_f^{Occ}	D_f^{BOS}	D_f^{Max}	D_f^{C3-C5}	D_f^{C5-C7}	D_f^{Hyoid}	D_f^{Larynx}	$D_f^{Mandible}$	$D_f^{PG>26\ Gy}$	$D_f^{PG\leq 26\ Gy}$	$V_f^{PG>26\ Gy}$	$V_f^{PG\leq 26\ Gy}$	V_f^{neck}
1	**13** (X)			**15** (27)	**12** (20)	**13** (X)		**9** (14)	**11** (13) **9** (14)		**8** (11) **9** (15)		
2			**9** (X)	**NA**	**NA**	**NA**	**NA**	**15** (X)		**9** (16) **9** (14)		**11** (21) **10** (22)	**10** (15) ↑
3						**16** (X)		**18** (X)	**19** (25)		**10** (15) **8** (12)		**12** (X) ↓
4				**18** (29)	**19** (30)		**18** (X)				**17** (X) **15** (X)		**17** (26) ↑
5	**13** (X)	**20** (X)	**11** (X)	**15** (X)	**NA**		**NA**	**11** (X)	**14 (**21) **20** (X)		**9** (18) **9 (**17)		**10** (23) ↓
6	**19** (X)	**20** (X)	**NA**			**17** (X)		**17** (25)	**14** (20)		**8** (19)	**9** (16)	
7			**NA**		**NA**				**18** (29)	**7** (19)	**11** (16)	**12** (21)	**NA**
8		**10** (X)	**12** (X)	**11** (X)	**NA**		**20** (X)	**8** (14)	**13** (16) **15** (18)		**20** (25) **11** (17)		**9** (13) ↓
9	**12** (X)	**9** (X)			**NA**	**11** (26)	**20** (X)		**13** (15)	**18** (23)	**12** (19)		
10			**15** (X)	**13** (X)	**18** (X)	**14** (X)	**NA**	**17** (X)	**18** (24)		**9** (14)	**20** (24)	**14** (27) ↓
11		**9** (X)	**8** (X)	**10** (X)	**14** (X)	**12** (14)	**14** (X)		**11** (28)		**12** (20)		**18** (28) ↓
12		**20** (X)		**14** (23)	**13** (20)	**16** (X)			**12** (17)	**9** (X)	**8** (12)	**12** (18)	**15** (22) ↓
Nr of cases	**0/4**	**0/6**	**0/5**	**3/7**	**3/5**	**2/7**	**0/4**	**3/7**	**12/13**	**4/5**	**14/16**	**6/6**	**7/8**
Efficiency	**0%**	**0%**	**0%**	**43%**	**60%**	**29%**	**0%**	**43%**	**92%**	**80%**	**88%**	**100%**	**88%**
95% CI	**0-13%**	**0-8%**	**0-10%**	**0-87%**	**7-100%**	**0 - 69%**	**0 -13%**	**0% -87%**	**74-100%**	**35-100%**	**68-100%**	**92-100%**	**58-100%**

A.5: Individual patient 90% level of confidence clinical limit results

Table A.5.1: Fraction number when exponentially weighted moving average (EWMA) statistics exceeded the 2 mm clinical limit during the first 4 weeks of treatment shown in boldface. Fraction number when the accuracy interval 2 mm clinical limit shown in parentheses. EWMA parameters are established from deformable image registration (DIR) raw data. (No entry = deviations within 2 mm or exceeded control after treatment week 4; NA = not assessed; BOS = base of skull; Occ = occipital; Max = maxilla; Mand = mandible; X = accuracy interval did not exceed the process limit 90% level of confidence).

Patient number	D_f^{Occ}	D_f^{BOS}	D_f^{Max}	D_f^{C3-C5}	D_f^{C5-C7}	D_f^{Hyoid}	D_f^{Larynx}	$D_f^{Mandible}$
1					**19** (21)	**1** (1)	**1** (1)	**10** (14)
2	**1** (1)	**1** (X)	**1** (1)	**NA**	**NA**	**NA**	**NA**	**1** (1)
3	**1** (1)					**1** (3)	**1** (1)	
4	**1** (X)				**1** (18)	**1** (17)	**1** (9)	
5			**11** (X)		**NA**	**1** (13)	**NA**	
6			**NA**			**1** (17)	**1** (17)	
7			**NA**		**NA**	**1** (1)	**1** (1)	**1** (X)
8					**NA**	**1** (1)	**1** (1)	
9	**1** (12)				**NA**		**1** (19)	**1** (18)
10						**1** (1)	**NA**	
11						**1** (1)	**1** (1)	
12					**1** (1)		**1** (X)	**9** (X)
Nr of cases	**3/4**	**0/1**	**1/2**	**0/0**	**3/3**	**9/9**	**8/9**	**3/5**
Efficiency	**75%**	**0%**	**50%**	**NA**	**100%**	**100%**	**89%**	**60%**
95% CI	**20 -100%**	**0 -50%**	**0 -100%**	**NA**	**83 -100%**	**94 -100%**	**63 -100%**	**7 -100%**

A.6: Differences between E0 values of SPC charts based on DIR or MMR

Table A.6.1: Deformation (mm) of the exponentially weighted moving average (EWMA) reference period E_0, acquired with deformable image registration (DIR). (NA = not assessed; BOS = base of skull; Occ = occipital bone; Max = maxilla).

Patient number	E_0^{Occ}	E_0^{BOS}	E_0^{Max}	E_0^{C3-C5}	E_0^{C5-C7}	E_0^{Hyoid}	E_0^{Larynx}	$E_0^{Mandible}$
1	1.0	1.1	1.0	1.2	1.7	5.1	3.9	1.9
2	2.5	2.0	2.6	**NA**	**NA**	**NA**	**NA**	4.0
3	2.7	1.3	1.3	1.2	1.6	3.1	3.4	1.1
4	2.2	0.9	1.7	1.3	2.2	2.3	2.5	1.3
5	1.2	1.9	2.0	1.2	**NA**	2.5	**NA**	1.2
6	1.7	1.4	**NA**	0.9	1.0	2.1	2.7	1.3
7	1.7	0.9	**NA**	1.2	**NA**	4.1	5.9	2.5
8	1.7	0.7	1.2	1.1	**NA**	3.4	2.8	1.4
9	2.2	1.5	1.2	1.4	**NA**	1.5	2.1	2.1
10	1.5	0.8	0.8	0.8	1.0	5.7	**NA**	0.9
11	1.2	1.5	1.4	1.1	1.6	3.0	3.9	1.8
12	1.7	1.0	0.9	1.6	2.5	1.5	2.2	1.9
Mean	**1.8**	**1.2**	**1.4**	**1.2**	**1.7**	**3.1**	**3.3**	**1.8**

Table A.6.2: Deformation (mm) of the exponentially weighted moving average (EWMA) reference period E_0, acquired with the manual method of registration (MMR). (NA = not assessed; BOS = base of skull; Occ = occipital bone; Max = maxilla).

Patient number	E_0^{Occ}	E_0^{BOS}	E_0^{Max}	E_0^{C3-C5}	E_0^{C5-C7}	E_0^{Hyoid}	E_0^{Larynx}	$E_0^{Mandible}$
1	1.2	1.1	2.3	1.4	2.6	4.6	4.8	3.5
2	2.5	1.9	3.8	**NA**	**NA**	**NA**	**NA**	5.3
3	1.3	1.1	2.2	1.5	2.3	3.5	2.3	1.5
4	1.1	0.7	2.0	1.2	1.8	2.2	1.9	1.4
5	1.8	1.4	1.7	1.6	**NA**	2.0	**NA**	1.9
6	0.6	0.8	**NA**	1.0	1.3	2.4	3.0	1.6
7	2.0	1.2	**NA**	1.0	**NA**	3.4	4.6	2.6
8	1.6	0.6	1.5	1.2	**NA**	3.9	3.0	2.2
9	1.3	1.5	2.1	1.5	**NA**	2.0	2.1	2.8
10	1.0	0.8	1.0	0.9	1.2	6.7	**NA**	1.3
11	1.4	1.1	1.7	1.6	2.0	2.0	2.6	2.1
12	1.1	1.3	1.6	1.5	2.0	1.8	1.9	2.1
Mean	**1.4**	**1.1**	**2.0**	**1.3**	**1.9**	**3.1**	**2.9**	**2.4**

*Table A.6.3: Differences in deformation (mm) of the exponentially weighted moving average (EWMA) reference periods E_0, acquired with deformable image registration (DIR) and the manual method of registration (MMR). (NA = not assessed; BOS = base of skull; Occ = occipital bone; Max = maxilla. *indicates a significant difference at the 0.95% level of confidence).*

Patient number	E_0^{Occ}	E_0^{BOS}	E_0^{Max}	E_0^{C3-C5}	E_0^{C5-C7}	E_0^{Hyoid}	E_0^{Larynx}	$E_0^{Mandible}$
1	-0.2	0.0	-1.2	-0.2	-0.9	0.5	-0.9	-1.5
2	0.0	0.1	-1.2	**NA**	**NA**	**NA**	**NA**	-1.3
3	1.4	0.1	-0.9	-0.3	-0.7	-0.4	1.1	-0.4
4	1.1	0.3	-0.3	0.2	0.4	0.1	0.6	-0.1
5	-0.6	0.5	0.2	-0.5	**NA**	0.5	**NA**	-0.7
6	1.1	0.6	**NA**	-0.1	-0.2	-0.3	-0.3	-0.4
7	-0.3	-0.3	**NA**	0.2	**NA**	0.7	1.3	-0.1
8	0.0	0.1	-0.2	-0.1	**NA**	-0.5	-0.2	-0.8
9	0.9	-0.1	-0.9	-0.1	**NA**	-0.4	0.0	-0.6
10	0.5	0.0	-0.1	-0.1	-0.2	-1.1	**NA**	-0.4
11	-0.2	0.4	-0.3	-0.6	-0.3	0.9	1.3	-0.3
12	0.5	-0.3	-0.7	0.0	0.5	-0.3	0.3	-0.2
Mean difference	**0.4**	**0.1**	**-0.6***	**-0.1**	**-0.2**	**0.0**	**0.3**	**-0.6***

A.7: Average parotid shift and volume EWMA

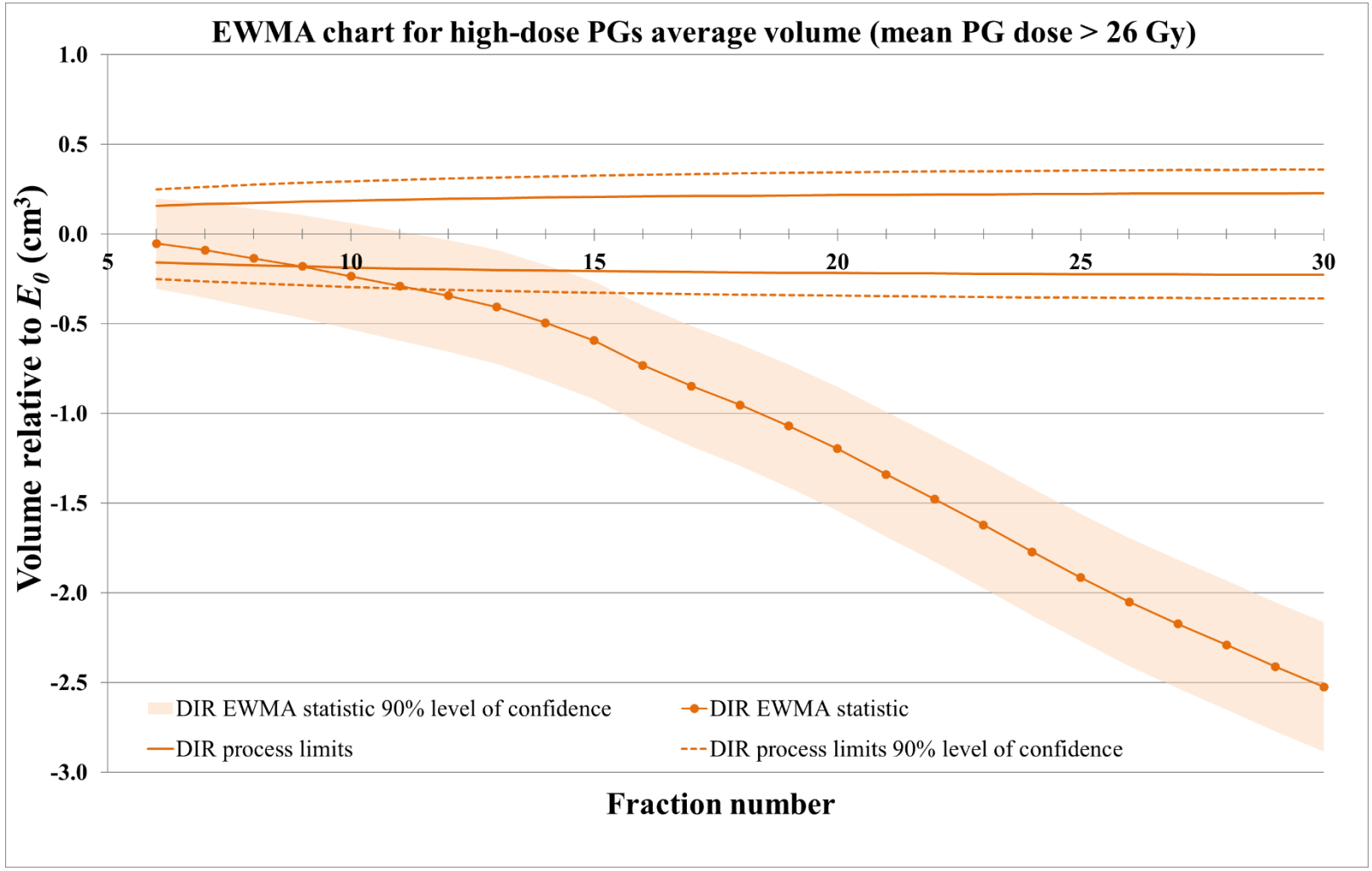


Figure A.7.1: Exponentially weighted moving average (EWMA) chart of average volume change of the high-dose parotid gland (PG) group.

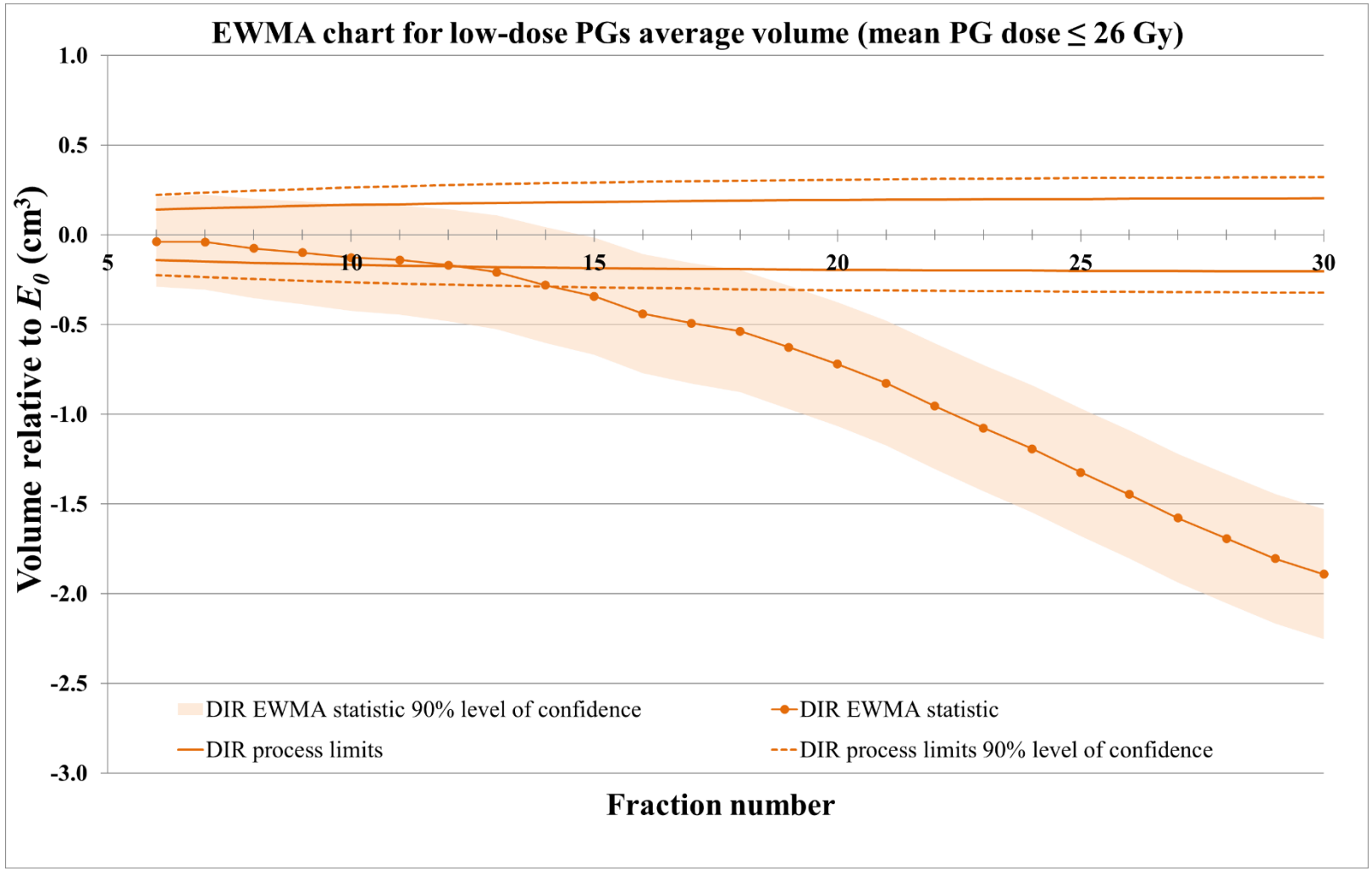


Figure A.7.2: Exponentially weighted moving average (EWMA) chart of average volume change of the low-dose parotid gland (PG) group.

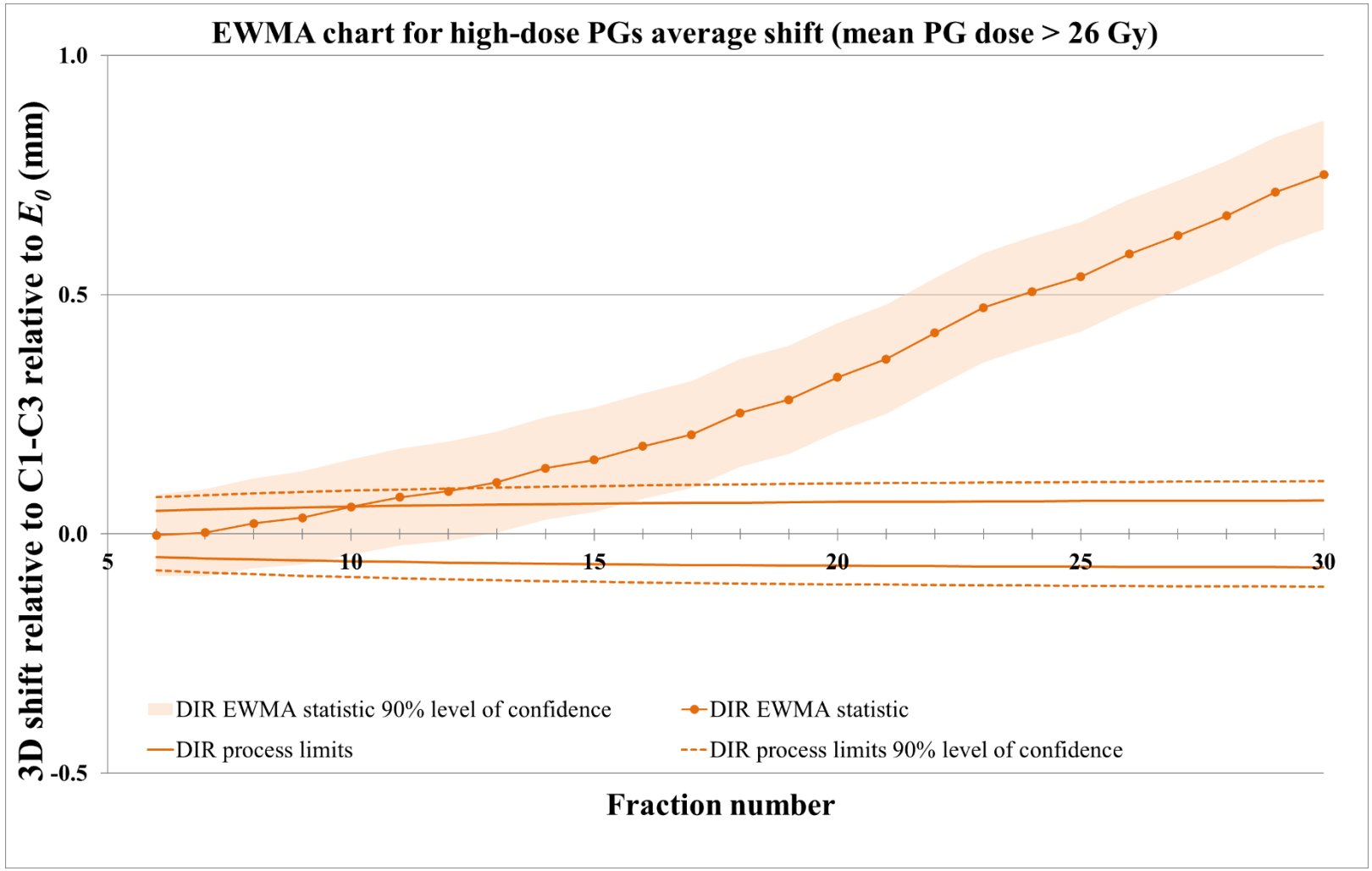


Figure A.7.3: Exponentially weighted moving average (EWMA) chart of average shift of the high-dose parotid gland (PG) group.

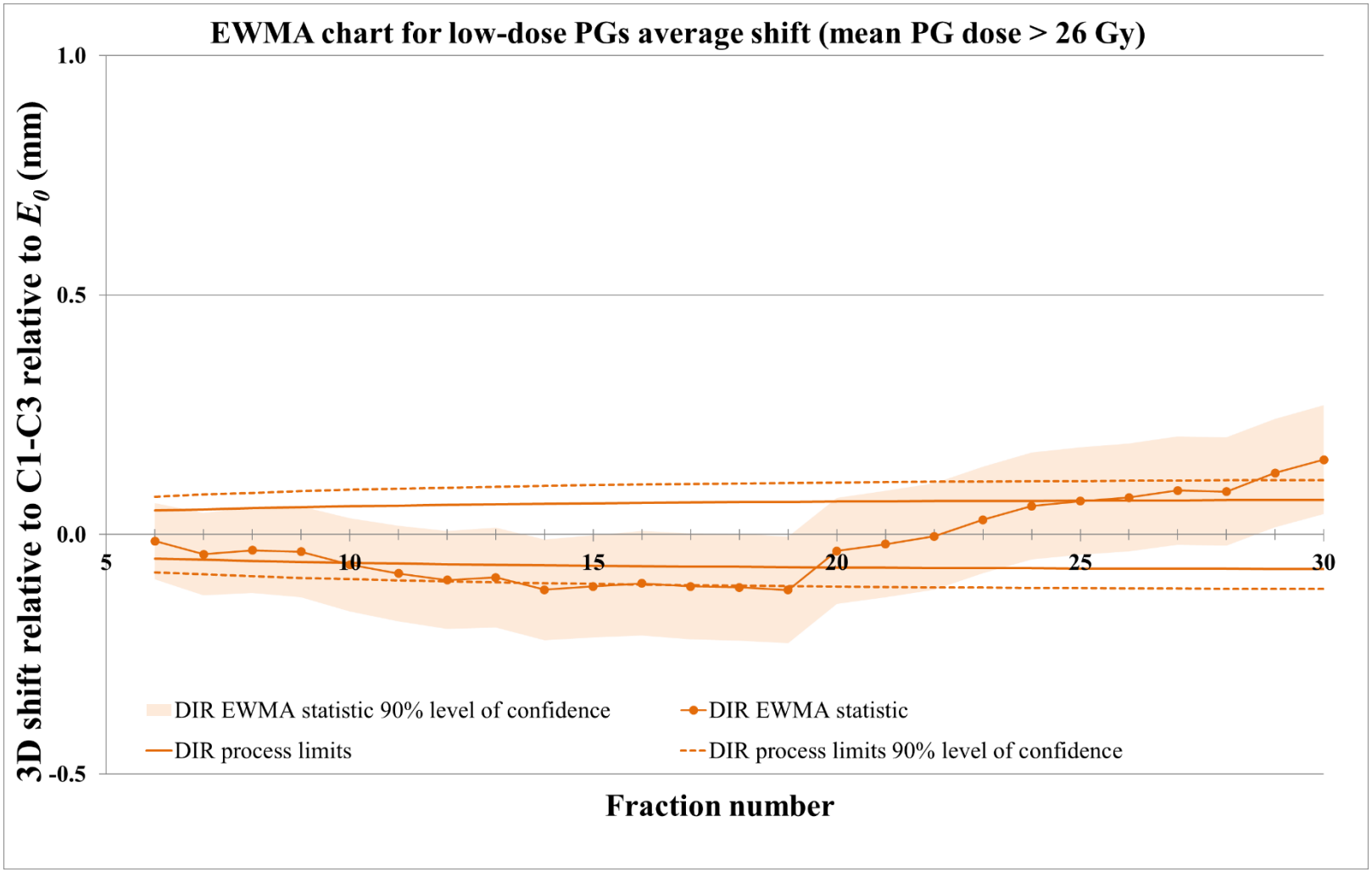


Figure A.7.4: Exponentially weighted moving average (EWMA) chart of average shift of the low-dose parotid gland (PG) group.

A.8: Radiation oncologist PG review

Table A.8.1: Summary of Dice similarity coefficient (DSC) and Hausdorff distance (HD) metrics evaluating the deformed parotid gland (PG) contours. Deformed PG contours were evaluated against independent radiation oncologist re-contours.

DSC		
	Mean	0.82
	SD	0.07
	Range	0.62 – 0.92
HD		
	Mean	0.98 cm
	SD	0.44 cm
	Range	0.40 – 2.47 cm

Table A.8.2: Patient-, fraction- and parotid gland-specific Dice similarity coefficient (DSC) and Hausdorff distance (HD) metrics evaluating the deformed parotid gland (PG) contours. Deformed PG contours were evaluated against independent radiation oncologist re-contours.

Patient, CBCT fraction, parotid gland	***DSC***	***HD (cm)***
1, CBCT10, left parotid	*0.90*	*0.48*
1, CBCT20, left parotid	*0.87*	*1.02*
1, CBCT30, left parotid	*0.86*	*0.47*
2, CBCT10, left parotid	*0.78*	*1.03*
2, CBCT20, left parotid	*0.73*	*1.29*
2, CBCT30, left parotid	*0.74*	*1.21*
3, CBCT10, left parotid	*0.88*	*0.72*
3, CBCT20, left parotid	*0.87*	*0.44*
3, CBCT30, left parotid	*0.85*	*0.46*
4, CBCT10, left parotid	*0.87*	*0.76*
4, CBCT20, left parotid	*0.85*	*1.11*
4, CBCT30, left parotid	*0.84*	*1.05*
5, CBCT10, left parotid	*0.72*	*1.25*
5, CBCT20, left parotid	*0.70*	*1.44*
5, CBCT30, left parotid	*0.62*	*1.61*
6, CBCT10, left parotid	*0.81*	*0.81*
6, CBCT20, left parotid	*0.86*	*1.08*
6, CBCT30, left parotid	*0.74*	*1.65*
7, CBCT10, left parotid	*0.87*	*0.48*
7, CBCT20, left parotid	*0.79*	*1.02*
7, CBCT30, left parotid	*0.64*	*1.07*
8, CBCT10, left parotid	*0.85*	*1.11*
8, CBCT20, left parotid	*0.82*	*0.96*
8, CBCT30, left parotid	*0.77*	*1.17*
9, CBCT10, left parotid	*0.81*	*0.89*
9, CBCT20, left parotid	*0.82*	*0.99*
9, CBCT30, left parotid	*0.77*	*0.70*
10, CBCT10, left parotid	*0.92*	*0.41*
10, CBCT20, left parotid	*0.89*	*0.61*
10, CBCT30, left parotid	*0.90*	*0.41*
11, CBCT10, left parotid	*0.92*	*0.40*

11, CBCT20, left parotid	*0.91*	*0.46*
11, CBCT30, left parotid	*0.82*	*0.67*
12, CBCT10, left parotid	*0.88*	*1.28*
12, CBCT20, left parotid	*0.86*	*1.54*
12, CBCT30, left parotid	*0.83*	*0.90*
1, CBCT10, right parotid	*0.92*	*0.90*
1, CBCT20, right parotid	*0.89*	*0.62*
1, CBCT30, right parotid	*0.89*	*0.45*
2, CBCT10, right parotid	*0.77*	*0.87*
2, CBCT20, right parotid	*0.68*	*0.94*
2, CBCT30, right parotid	*0.70*	*0.88*
3, CBCT10, right parotid	*0.87*	*0.61*
3, CBCT20, right parotid	*0.87*	*0.59*
3, CBCT30, right parotid	*0.83*	*0.59*
4, CBCT10, right parotid	*0.84*	*0.80*
4, CBCT20, right parotid	*0.84*	*0.77*
4, CBCT30, right parotid	*0.84*	*1.06*
5, CBCT10, right parotid	*0.77*	*1.15*
5, CBCT20, right parotid	*0.69*	*1.86*
5, CBCT30, right parotid	*0.69*	*1.60*
6, CBCT10, right parotid	*0.79*	*1.67*
6, CBCT20, right parotid	*0.79*	*1.63*
6, CBCT30, right parotid	*0.79*	*1.37*
7, CBCT10, right parotid	*0.85*	*0.71*
7, CBCT20, right parotid	*0.82*	*0.69*
7, CBCT30, right parotid	*0.70*	*1.39*
8, CBCT10, right parotid	*0.83*	*1.27*
8, CBCT20, right parotid	*0.81*	*1.60*
8, CBCT30, right parotid	*0.80*	*1.26*
9, CBCT10, right parotid	*0.79*	*2.14*
9, CBCT20, right parotid	*0.82*	*2.47*
9, CBCT30, right parotid	*0.85*	*0.91*
10, CBCT10, right parotid	*0.91*	*0.42*
10, CBCT20, right parotid	*0.90*	*0.50*
10, CBCT30, right parotid	*0.90*	*0.47*
11, CBCT10, right parotid	*0.89*	*0.61*
11, CBCT20, right parotid	*0.89*	*0.42*
11, CBCT30, right parotid	*0.84*	*1.15*
12, CBCT10, right parotid	*0.83*	*0.98*
12, CBCT20, right parotid	*0.79*	*0.79*
12, CBCT30, right parotid	*0.80*	*1.19*

A.9: Planning protocol

Table A.9.1: Dose-volume objectives used in treatment plan optimization.

Dose-volume metrics
$D_{0.1cc}$ *= minimum dose to 0.1 cm^3 of the volume that receives the highest dose*
D_{max} *= maximum dose*
D_{mean} *= mean dose*
$D_{98\%}$ *= minimum dose to 98% of the volume*
$D_{99\%}$ *= minimum dose to 99% of the volume*
PRV *= planning risk volume. A 3 mm PRV was used*
PTV *= planning target volume*
CTV *= clinical target volume*
GTV *= gross tumor volume*

**Not included in the structure analysis of this paper due to inconsistent physician contouring between patients*

Structure	Dose-volume objective	Priority
Target structures		
PTVs (high-, intermediate- and low-risk)	$D_{98\%} \geq$ 95% prescribed dose	High
	D_{max} < 107% max prescribed dose	High
CTVs (high-, intermediate- and low-risk)	$D_{99\%} \geq$ 95% prescribed dose	High
	D_{max} < 107% max prescribed dose	High
GTV	$D_{99\%} \geq$ 95% prescribed dose	High
	D_{max} < 107% max prescribed dose	High
Organs at risk		
Brain stem PRV	$D_{0.1cc}$ < 54 Gy	High
Spinal cord PRV	$D_{0.1cc}$ < 50 Gy	High
Optic nerve/chiasm PRV	$D_{0.1cc}$ < 55 Gy	High
Parotid glands		
-Ipsilateral	D_{mean} < 26 Gy [25,59]	Medium
-Contralateral	D_{mean} < 26 Gy [25,59]	Medium
Submandibular glands		
-Ipsilateral	D_{mean} < 39 Gy [277]	Low
-Contralateral	D_{mean} < 39 Gy [277]	Low
Oral cavity*	D_{mean} < 35 Gy [2]	Medium
Pharyngeal constrictor muscles*	D_{mean} < 50 Gy [45]	Low
Larynx*	D_{mean} < 44 Gy [45]	Low

A.10: Symmetry and inverse consistency

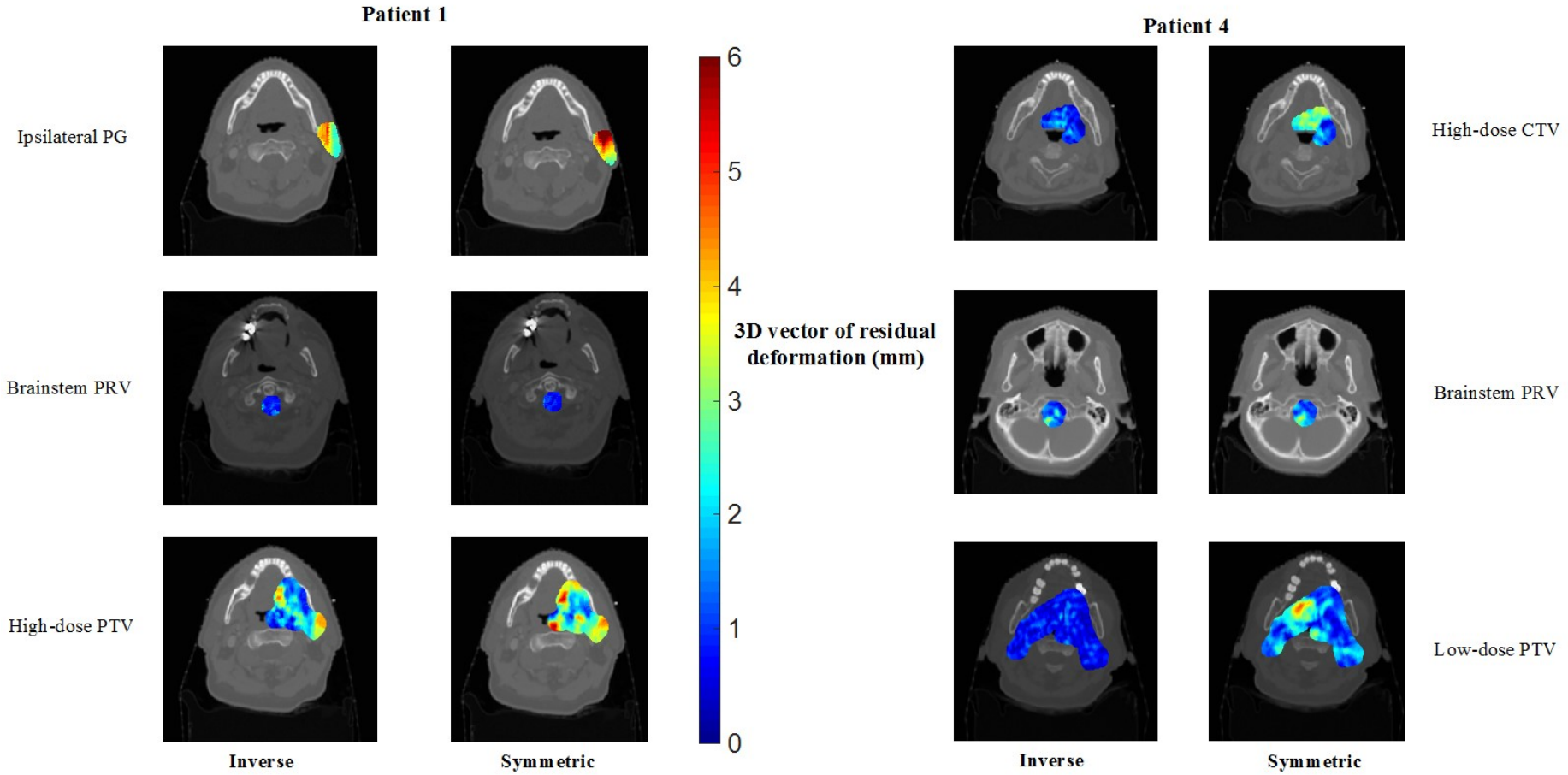


Figure A.10.1: 3D vector of residual deformations after applying the SmartAdapt (SA) acquired forward deformation vector field (DVF) and subsequent inverse (left columns) or symmetric (right columns) DVF registration. Structures of interest in head-and-neck radiotherapy (HNRT) are shown for patient 1 at fraction 29 and patient 4 at fraction 3 to exemplify the residual deformation errors. In general, the symmetric registrations had larger residual deformations relative to the inverse registrations. A notable inverse consistency error was present for patient 1's ipsilateral PG and high-dose PTV. PG = parotid gland. PRV = planning risk volume. PTV = planning target volume. CTV = clinical target volume.

A.11: IC rates

Table A.11.1: Structure- and fraction-specific inverse consistency (IC) rates [%] for patients 1 – 6. PTV = planning target volume. CTV = clinical target volume. PRV = planning risk volume. PG = parotid gland. SMG = submandibular gland.

Patient	***1***			***2***			***3***			***4***			***5***			***6***		
fraction	***3***	***16***	***29***	***3***	***16***	***29***	***3***	***16***	***29***	***3***	***16***	***29***	***3***	***16***	***29***	***3***	***16***	***29***
High-dose PTV	98.0	87.6	85.6				100	100	100	100	100	99.9	100	100	100	100	100	100
Intermediate-dose PTV																		
Low-dose PTV	97.5	94.1	94.0	100	100	99.7	100	100	100	100	99.9	99.9	100	100	100	99.9	99.9	99.9
High-dose CTV	98.4	86.8	85.3				100	100	100	100	100	99.8	100	100	100	100	100	100
Intermediate-dose CTV																		
Low-dose CTV	98.0	91.4	91.2	100	100	100	100	100	100	100	100	99.9	100	100	100	100	100	100
Brainstem PRV	86.8	94.9	97.0	99.8	98.1	96.9	100	73.3	100	99.9	99.1	97.6	99.8	98.8	99.6	96.4	87.3	95.9
Spinal cord PRV	78.9	100	99.9	98.5	95.3	98.8	99.4	100	99.3	95.9	96.3	93.7	99.5	97.6	99.9	100	100	100
Brainstem	64.2	84.6	93.7	99.8	98.4	98.9	100.0	64.2	100.0	98.4	93.9	90.8	99.7	97.8	99.3	98.1	83.2	94.1
Spinal cord	71.9	100.0	99.7	94.6	90.5	96.2	99.2	100.0	98.2	94.7	96.1	91.6	99.8	96.2	100.0	98.9	98.1	96.5
Ipsilateral PG	99.4	42.5	26.3	100	99.3	100	100	100	100	100	100	100	100	100	100	99.7	100	100
Contralateral PG	94.6	96.0	86.2	99.9	99.8	99.9	100	100	100	100	100	100	100	100	100	100	100	100
Ipsilateral SMG	99.8	95.9	98.9	100	100	100	100	100	100	100	100	100	100	100	100			
Contralateral SMG	100	99.7	100	100	100	97.6	100	100	100	100	100	100	100	100	100	100	100	100

Table A.11.2: Structure- and fraction-specific inverse consistency (IC) rates [%] for patients 7 – 12. PTV = planning target volume. CTV = clinical target volume. PRV = planning risk volume. PG = parotid gland. SMG = submandibular gland.

Patient	***7***			***8***			***9***			***10***			***11***			***12***		
fraction	***3***	***16***	***29***	***3***	***16***	***29***	***3***	***16***	***29***	***3***	***16***	***29***	***3***	***16***	***29***	***3***	***16***	***29***
High-dose PTV	99.9	99.7	99.7	100	100	100	99.9	99.8	99.8	100	100	100	99.9	100	100	100	100	100
Intermediate-dose PTV	100	100	100	100	100	100	99.9	99.9	99.9	100	100	100						
Low-dose PTV	100	100	100	100	99.9	100	99.8	99.6	99.6	100	100	99.9	99.7	99.9	99.9	99.5	99.7	99.9
High-dose CTV	100	99.8	99.8	100	100	100	99.8	99.7	99.7	100	100	100	100	100	100	100	100	100
Intermediate-dose CTV	100	99.9	99.9	100	100	100	99.9	99.8	99.8	100	100	100						
Low-dose CTV	100	100	99.9	100	100	100	99.8	99.7	99.7	100	100	99.9	99.9	99.9	99.9	99.8	99.8	99.9
Brainstem PRV	96.8	96.3	99.6	100	100	100	98.1	98.9	99.8	99.9	99.6	95.9	95.4	95.6	95.6	98.6	97.1	99.9
Spinal cord PRV	98.1	99.9	99.9	99.9	99.8	99.9	98.3	98.4	97.3	100	99.4	99.5	95.4	92.8	94.1	100	100	100
Brainstem	95.5	92.6	98.3	100.0	99.9	100.0	93.9	95.3	97.9	100.0	100.0	98.7	83.6	80.7	96.1	97.8	97.1	100.0
Spinal cord	98.8	99.8	99.9	100.0	100.0	100.0	98.0	98.7	96.8	100.0	99.6	100.0	94.8	92.6	92.7	100.0	100.0	100.0
Ipsilateral PG	100	99.9	99.8	100	100	100	100	99.9	100	100	100	100	99.9	99.4	100	100	99.7	99.9
							100	100	100				99.4	99.9	99.5			
Contralateral PG	100	100	100	100	100	100				100	100	100				100	99.9	99.9
Ipsilateral SMG	100	100	100	100	100	100	100	100	100	100	100	100	99.9	99.9	100	100	100	100
							100	100	100				100	100	100			
Contralateral SMG	100	100	100	100	100	100				100	100	100				100	100	100

A.12: Q-Q probability plots

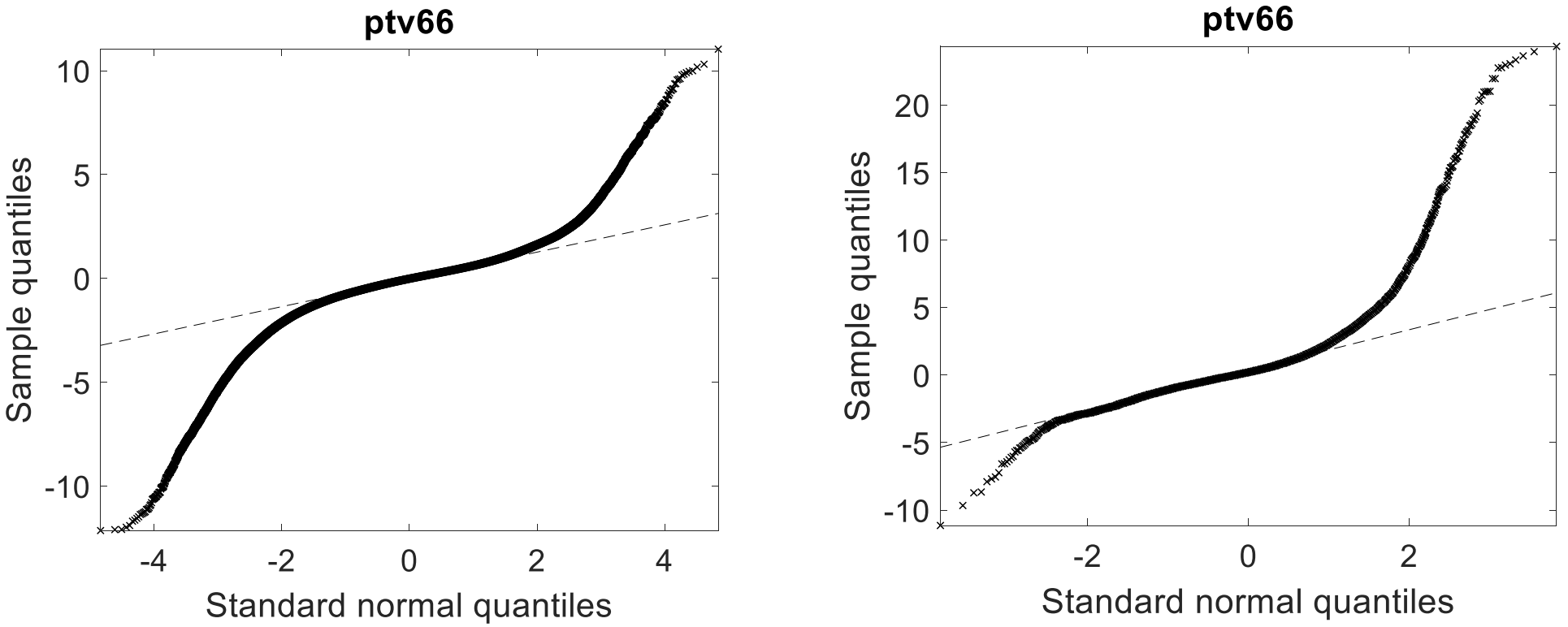


Figure A.12.1: Q-Q probability plots of the high-dose planning target volume (PTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

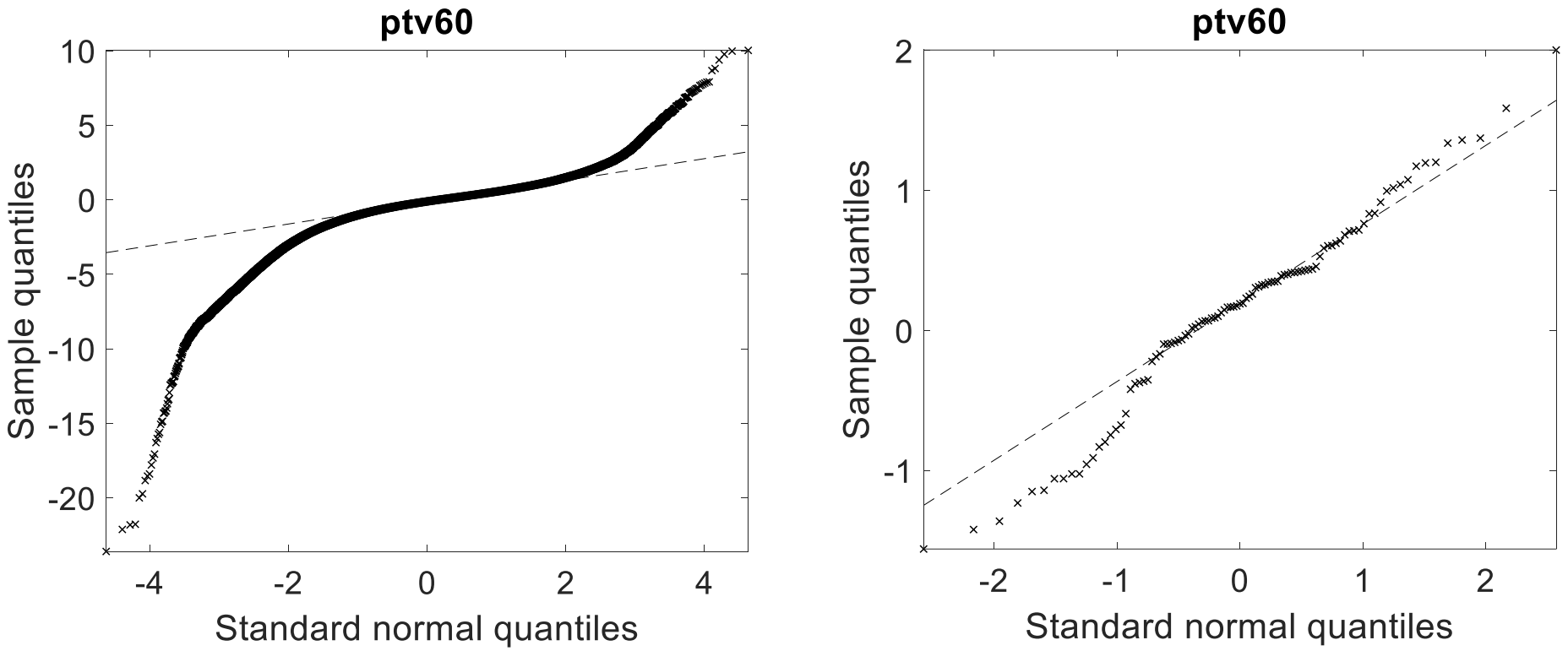


Figure A.12.2: Q-Q probability plots of the intermediate-dose planning target volume (PTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

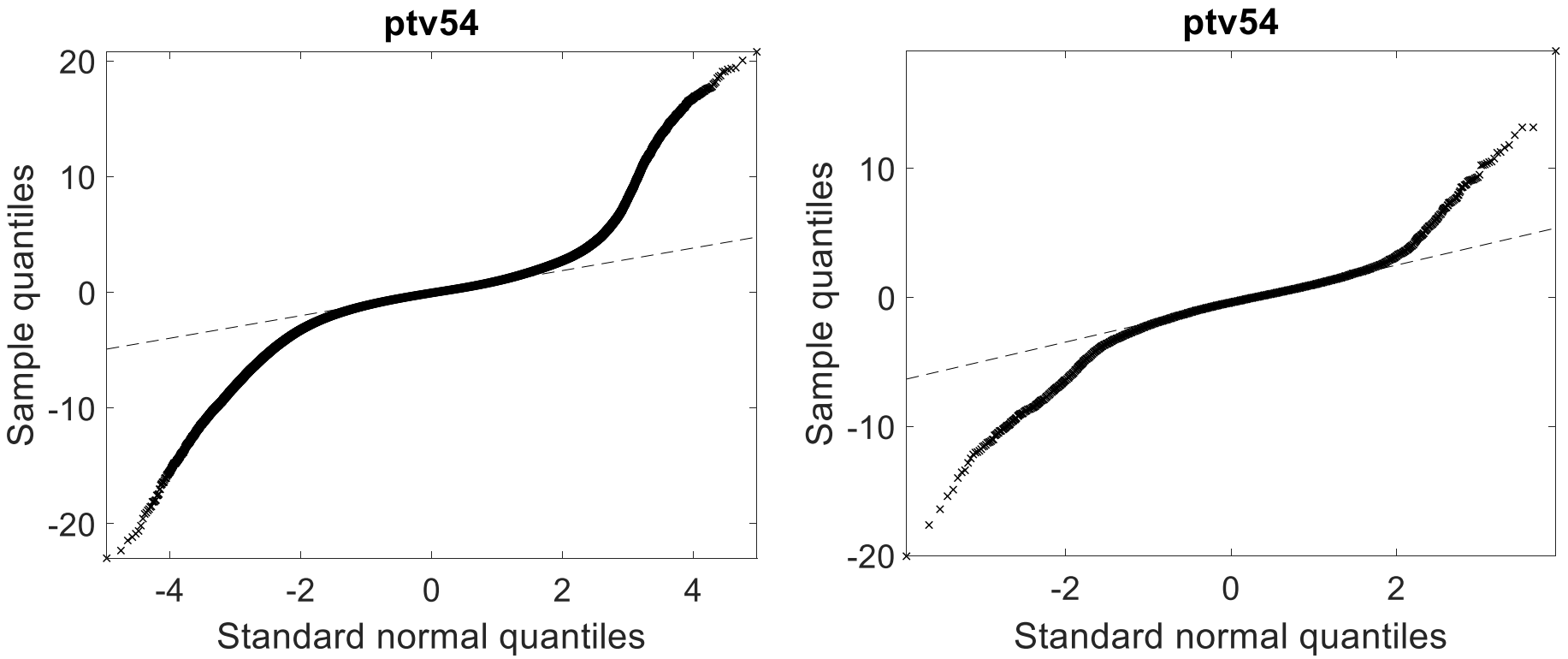


Figure A.12.3: Q-Q probability plots of the low-dose planning target volume (PTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

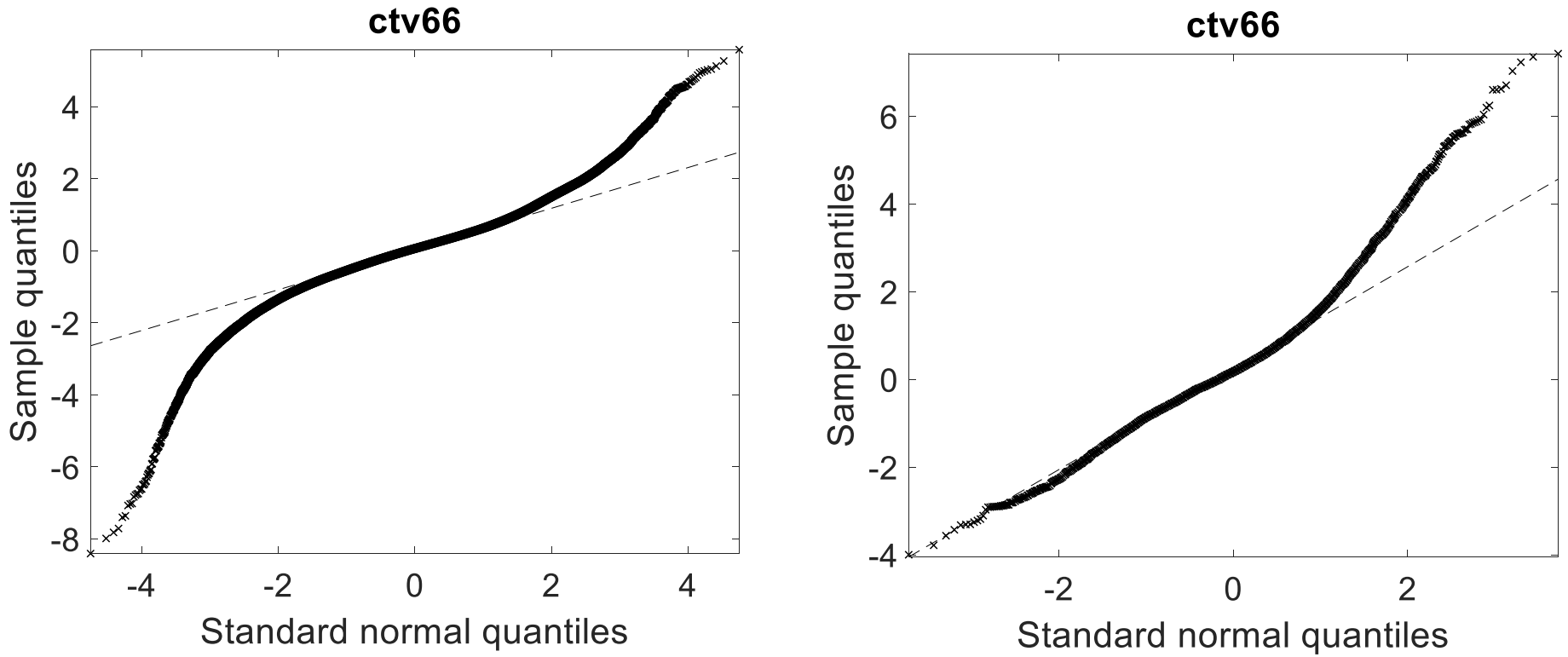


Figure A.12.4: Q-Q probability plots of the high-dose clinical target volume (CTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

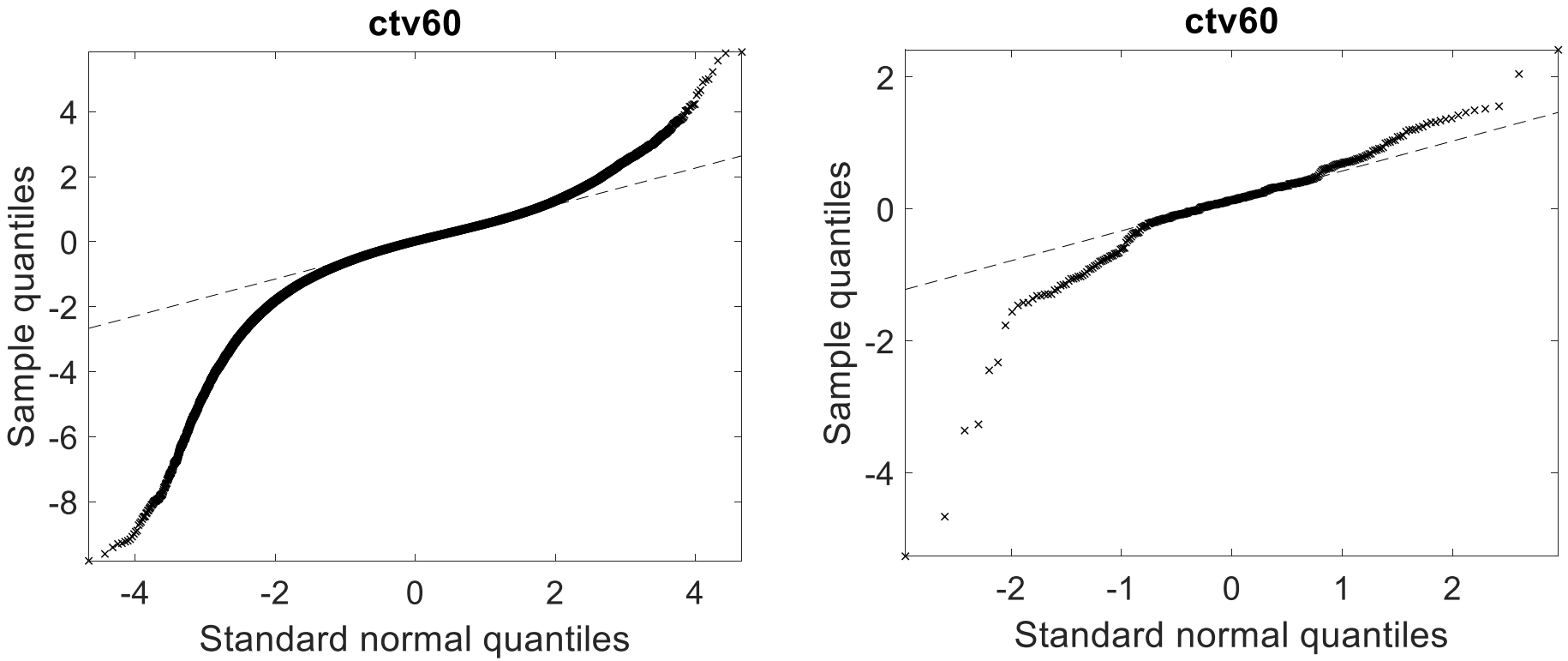


Figure A.12.5: Q-Q probability plots of the intermediate-dose clinical target volume (CTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

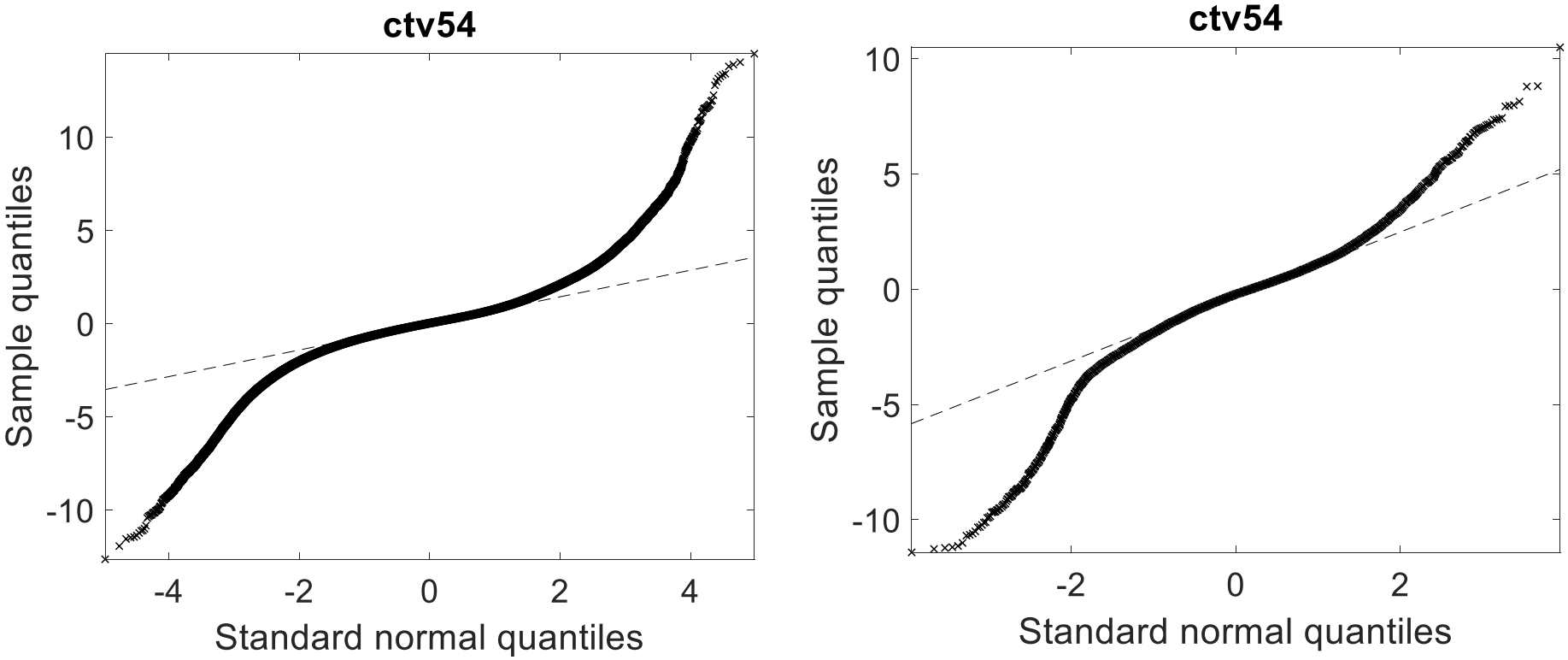


Figure A.12.6: Q-Q probability plots of the low-dose clinical target volume (CTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

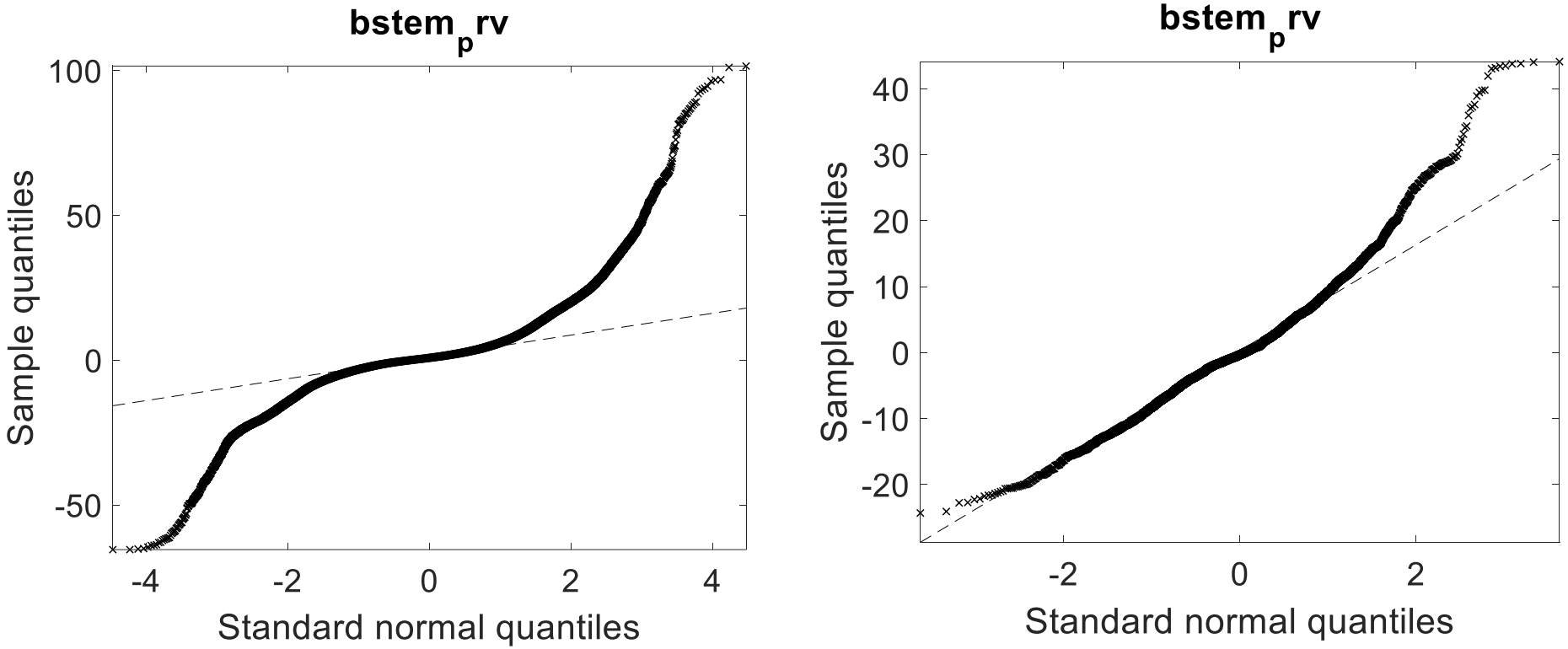


Figure A.12.7: Q-Q probability plots of the brainstem planning risk volume (PRV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

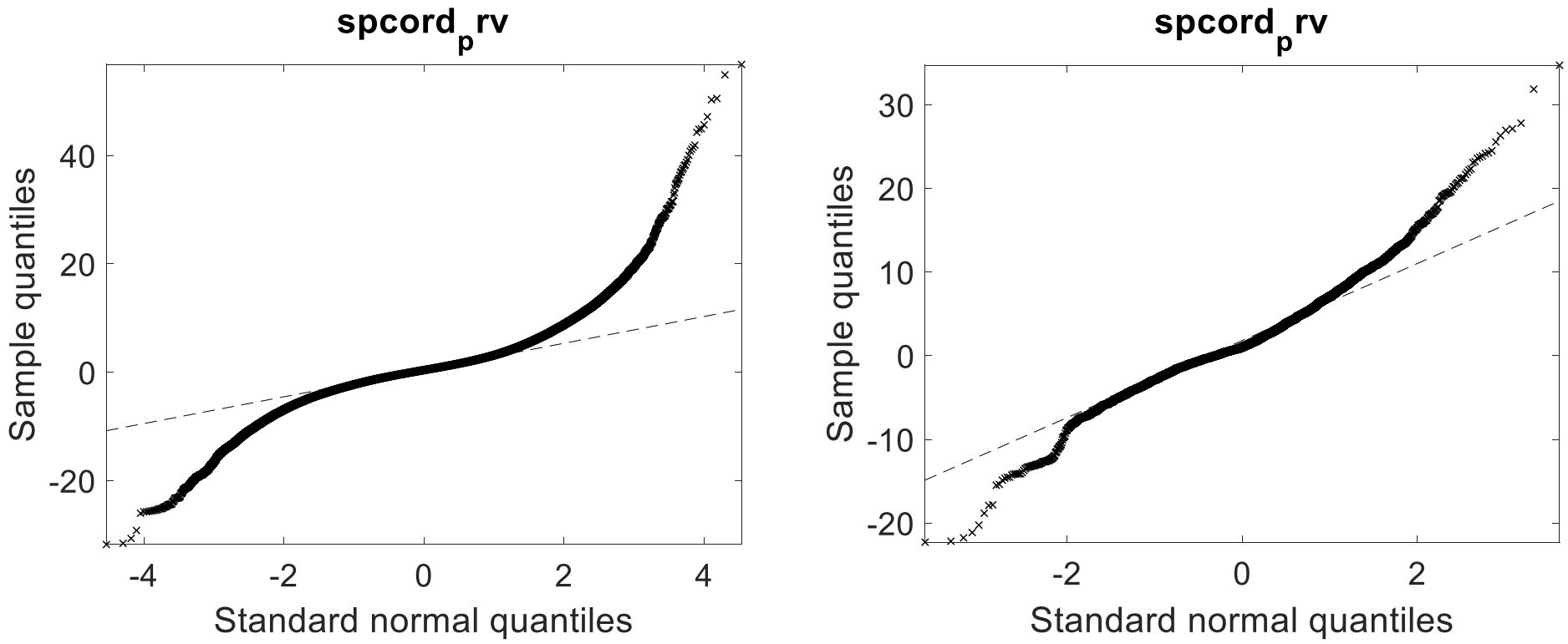


Figure A.12.8: Q-Q probability plots of the spinal cord planning risk volume (PRV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

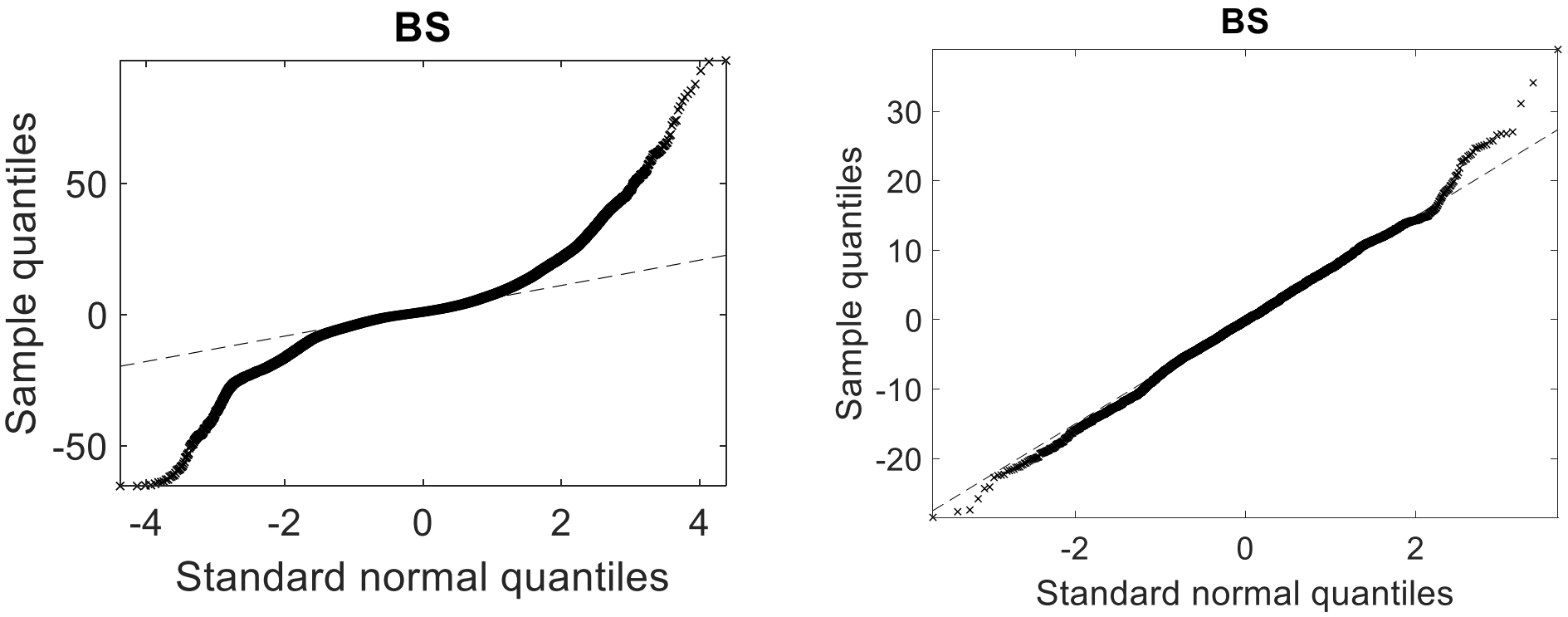


Figure A.12.9: Q-Q probability plots of the brainstem of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

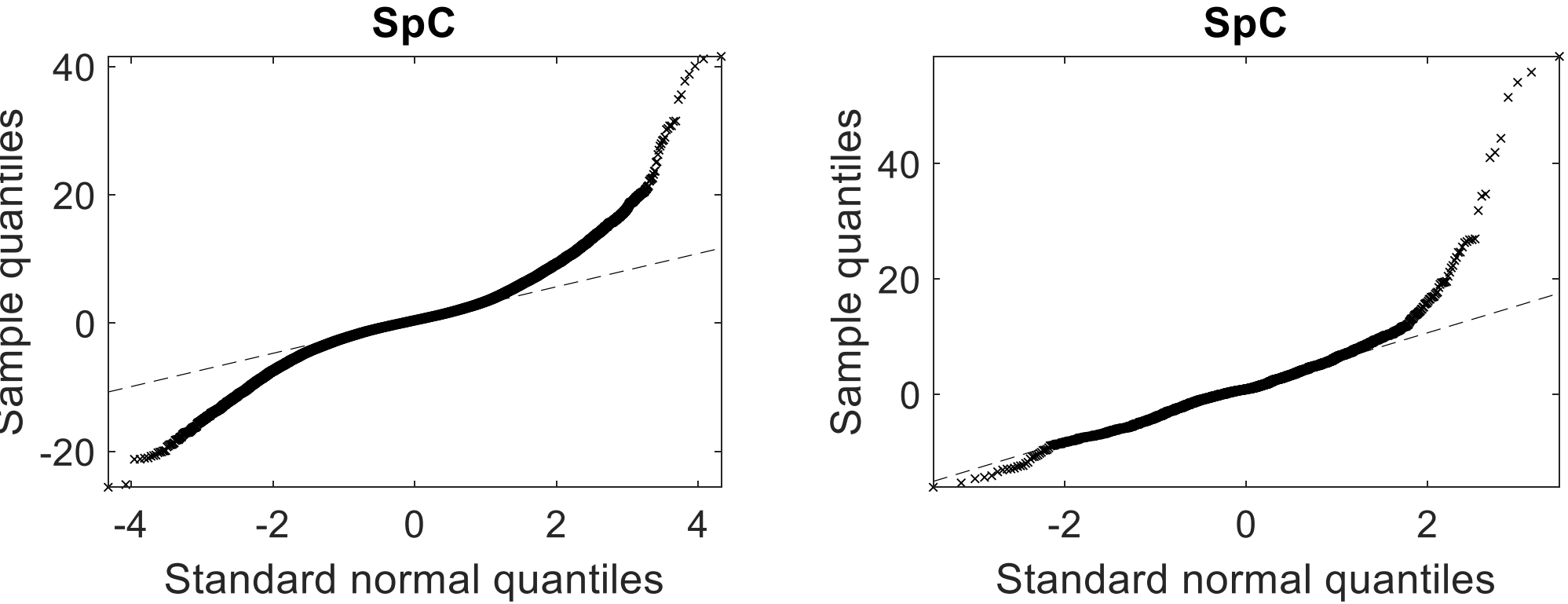


Figure A.12.10: Q-Q probability plots of the spinal cord of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

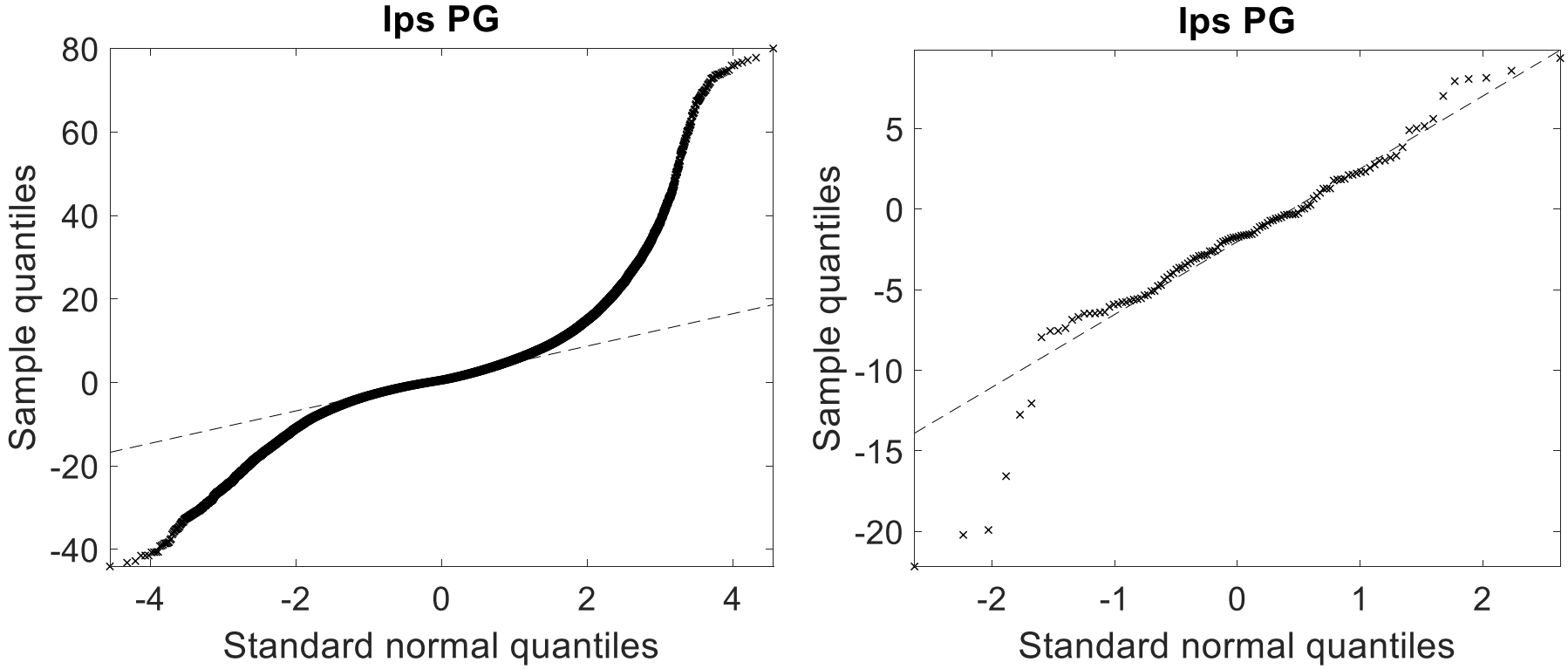


Figure A.12.11: Q-Q probability plots of the ipsilateral parotid gland (PG) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

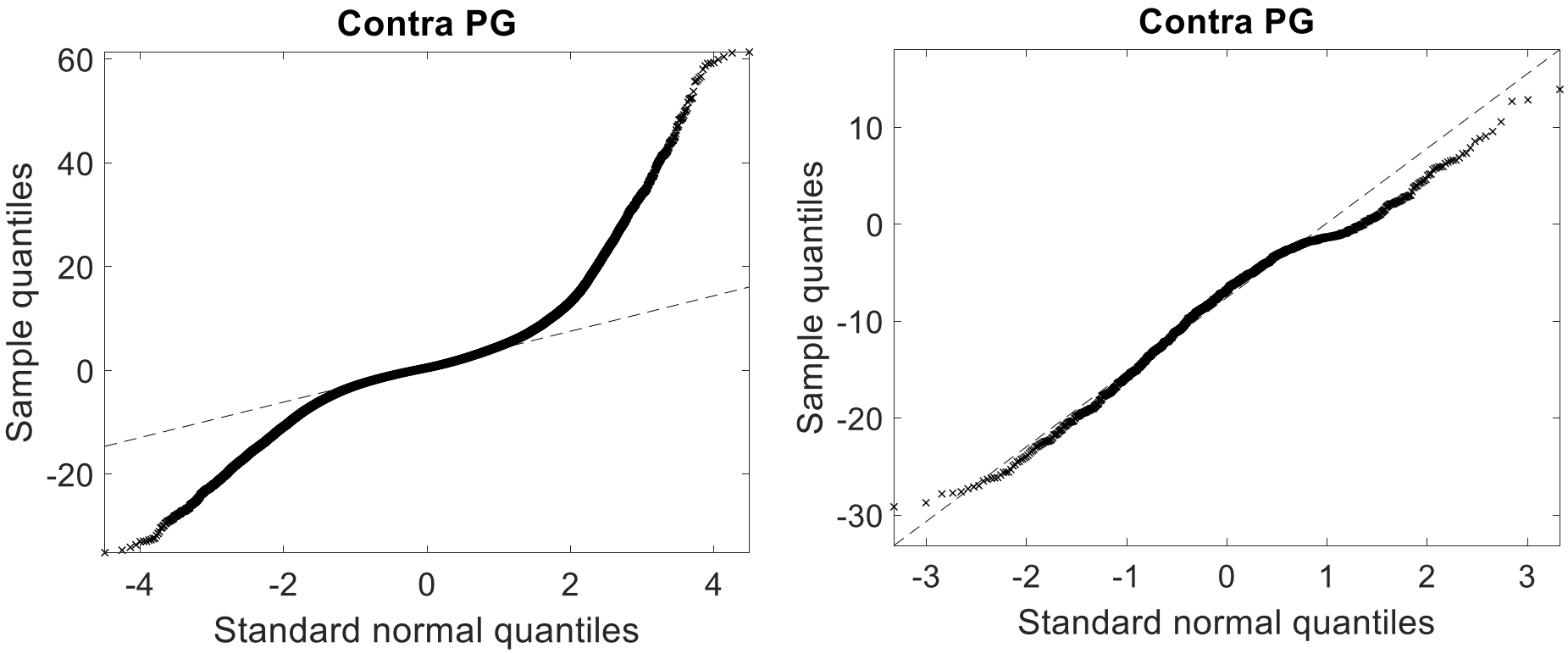


Figure A.12.12: Q-Q probability plots of the contralateral parotid gland (PG) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

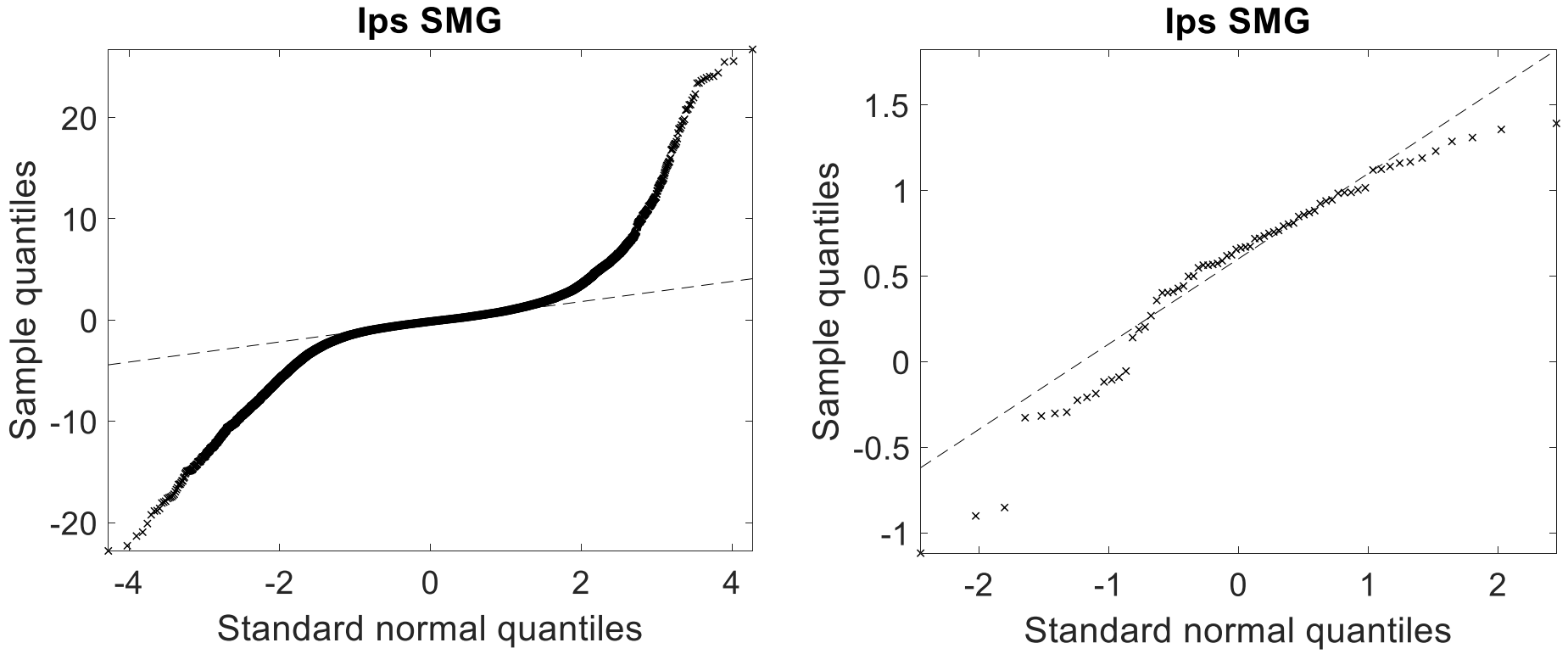


Figure A.12.13: Q-Q probability plots of the ipsilateral submandibular gland (SMG) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

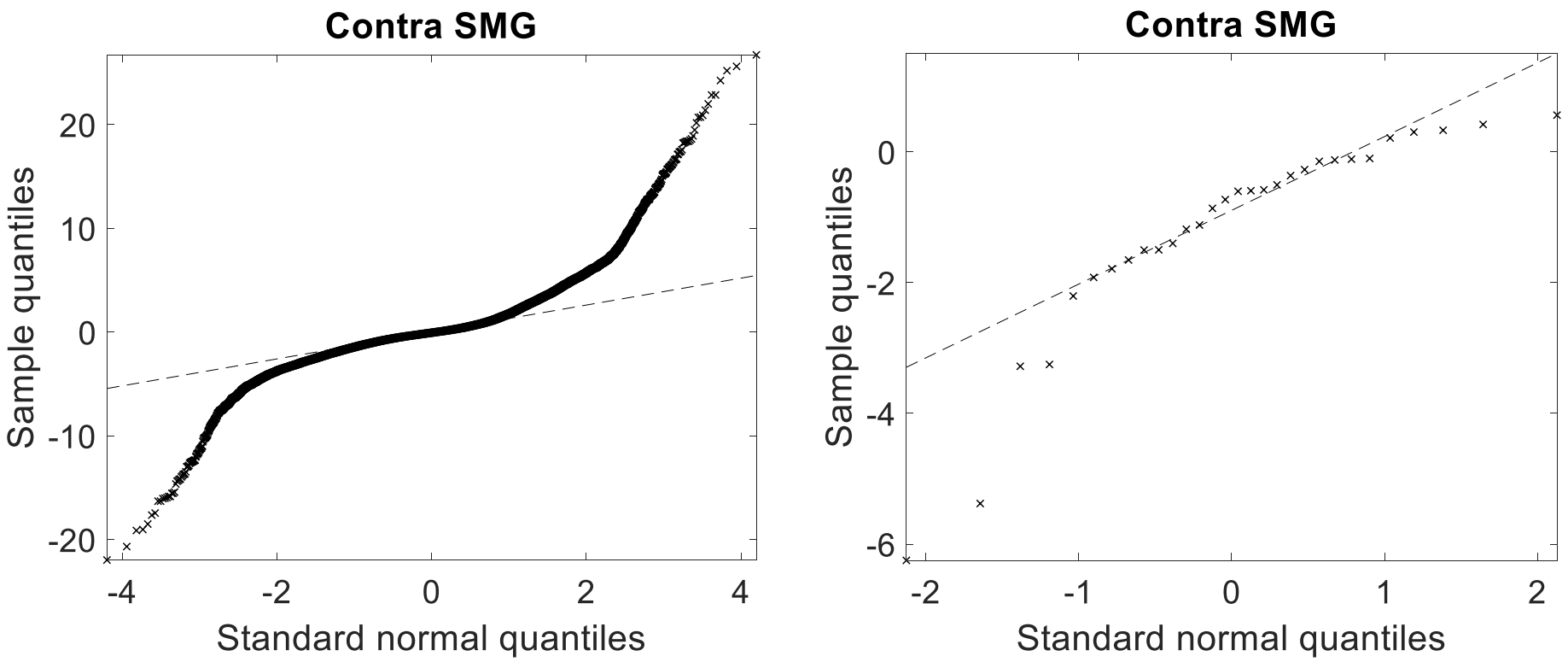


Figure A.12.14: Q-Q probability plots of the contralateral submandibular gland (SMG) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels.

A.13: Histograms of Δ^{tot}

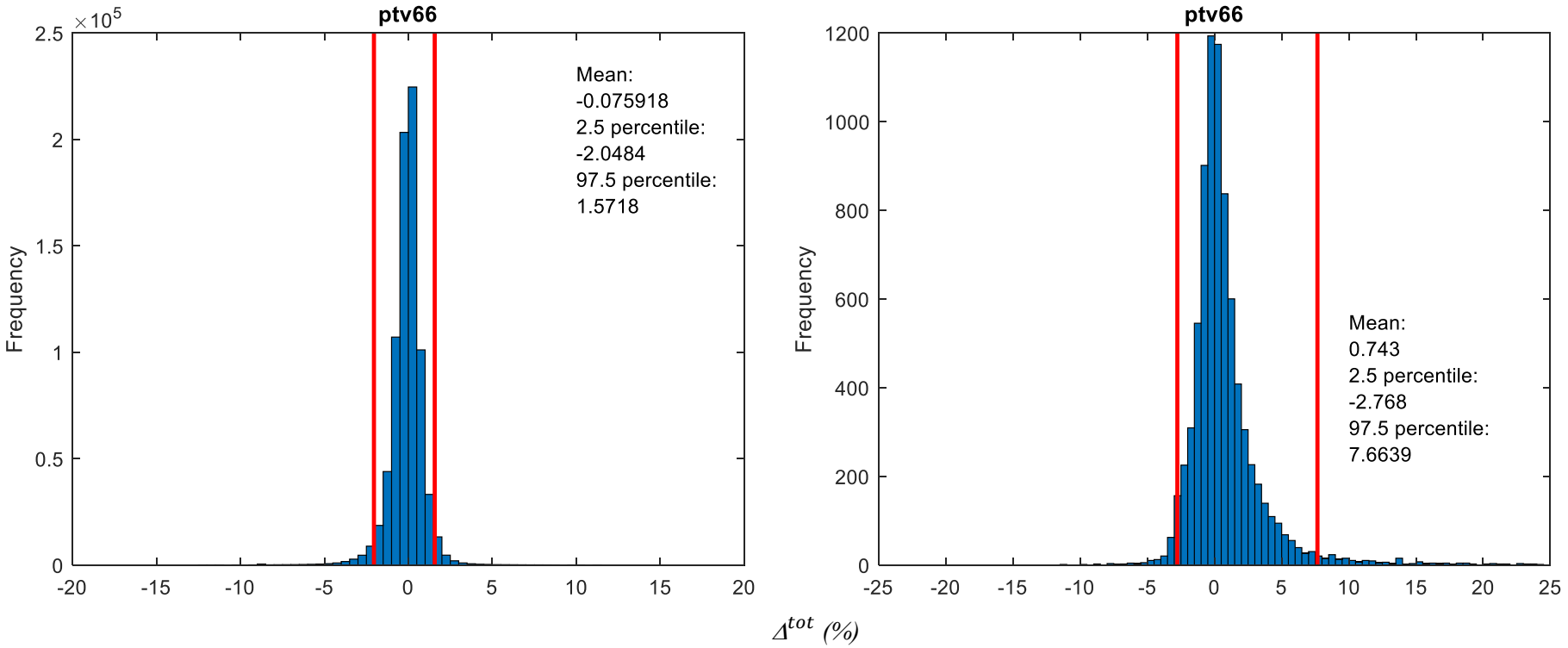


Figure A.13.1: Histogram showing the distribution of Δ^{tot} for the high-dose planning target volume (PTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

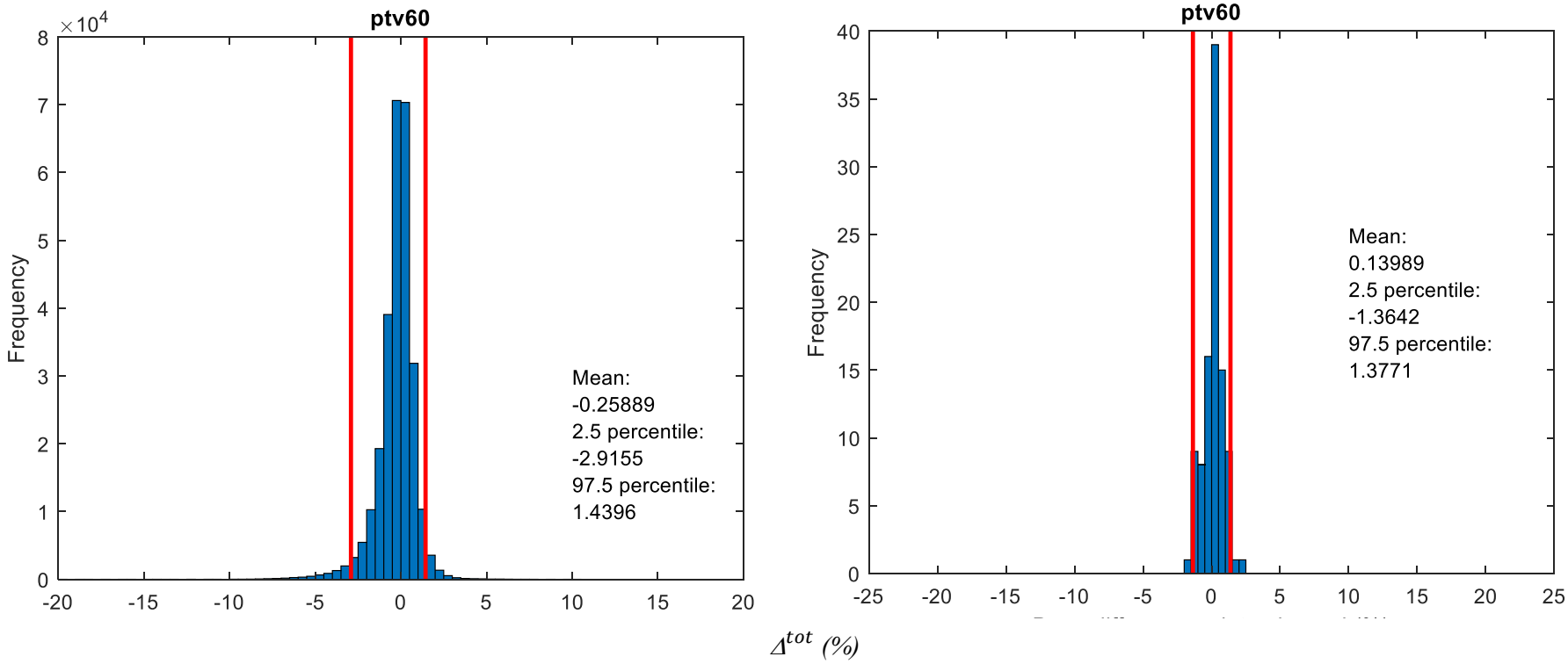


Figure A.13.2: Histogram showing the distribution of Δ^{tot} for the intermediate-dose planning target volume (PTV) of all patients for inverse consistent (left) inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

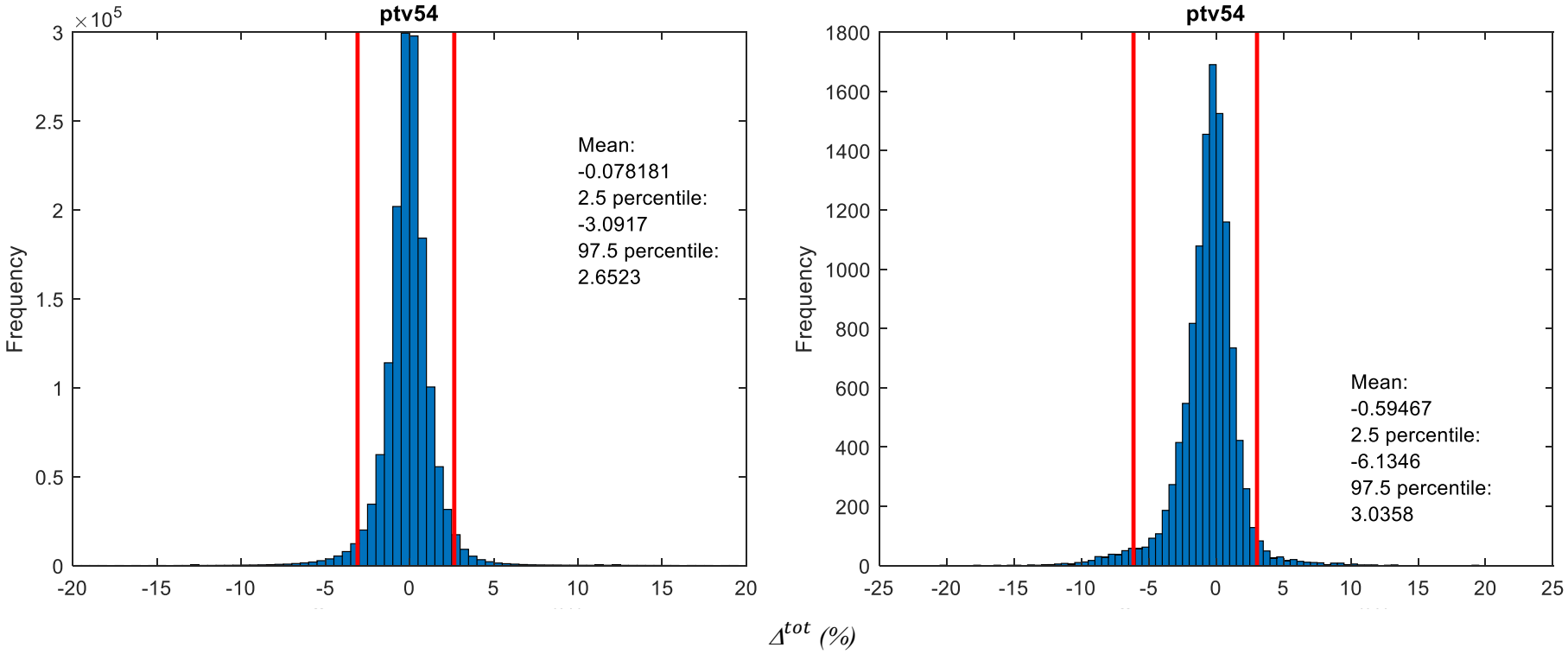


Figure A.13.3: Histogram showing the distribution of Δ^{tot} for the low-dose planning target volume (PTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

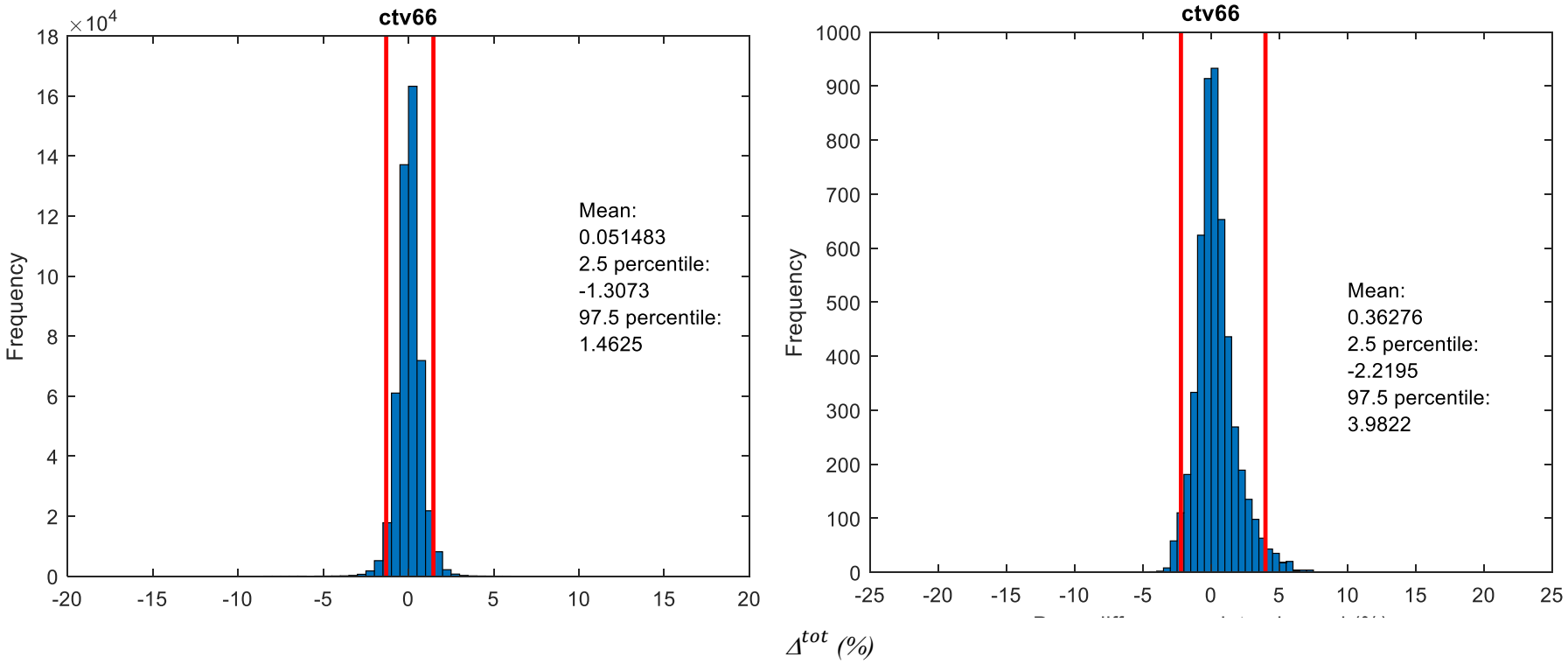


Figure A.13.4: Histogram showing the distribution of Δ^{tot} for the high-dose clinical target volume (CTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

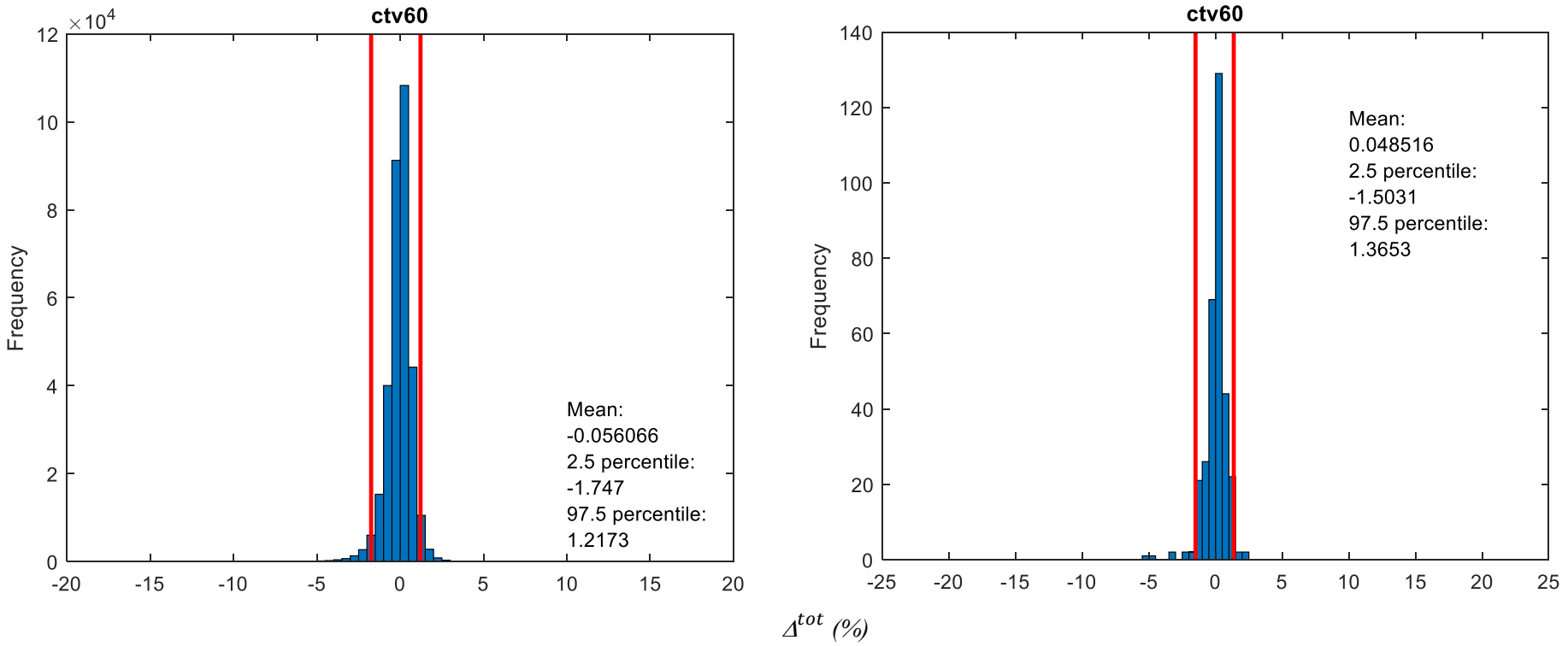


Figure A.13.5: Histogram showing the distribution of Δ^{tot} for the intermediate-dose clinical target volume (CTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

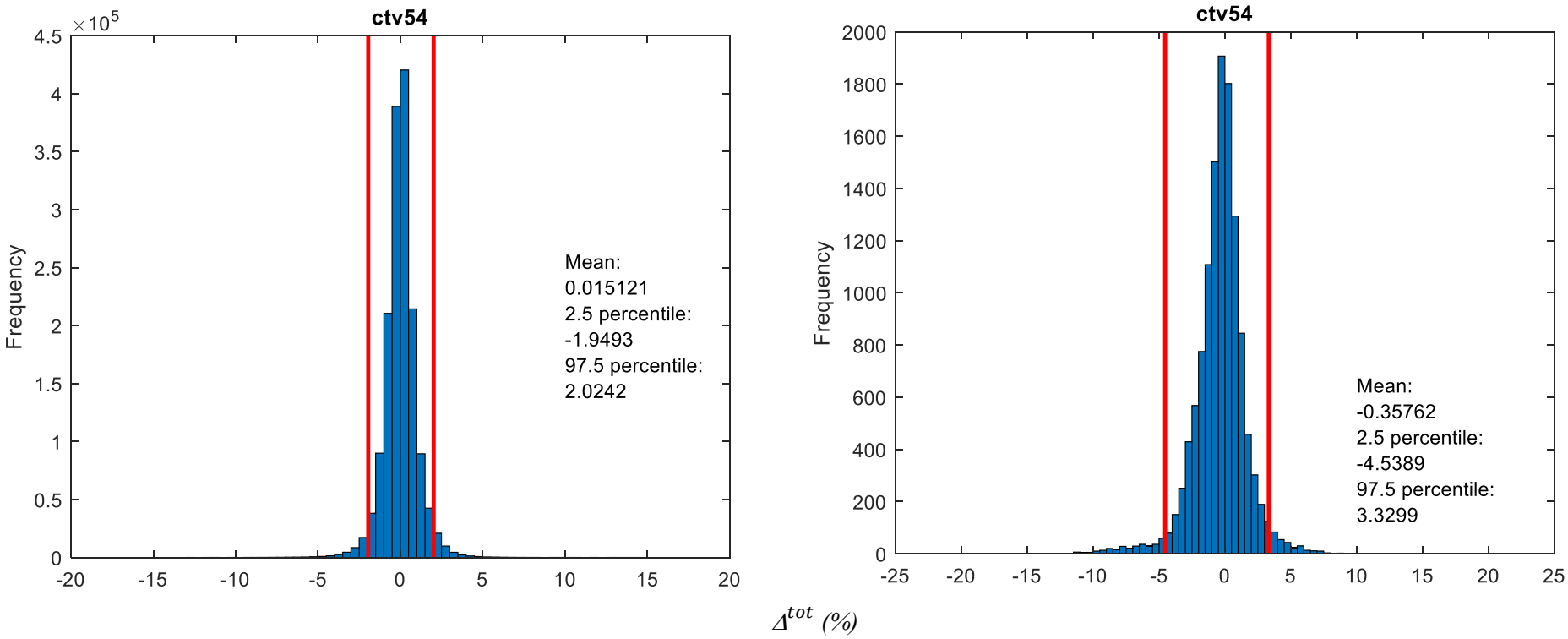


Figure A.13.6: Histogram showing the distribution of Δ^{tot} for the low-dose clinical target volume (CTV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

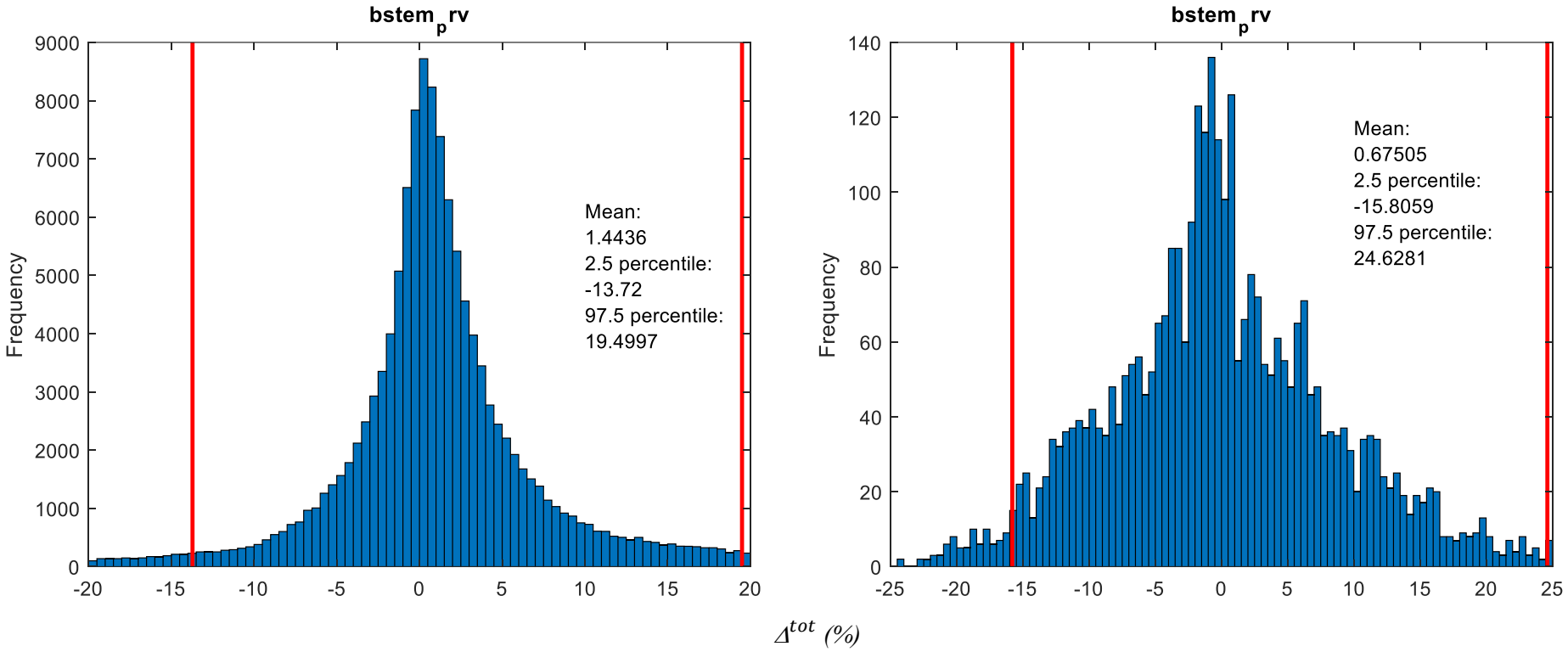


Figure A.13.7: Histogram showing the distribution of Δ^{tot} for the brainstem planning risk volume (PRV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

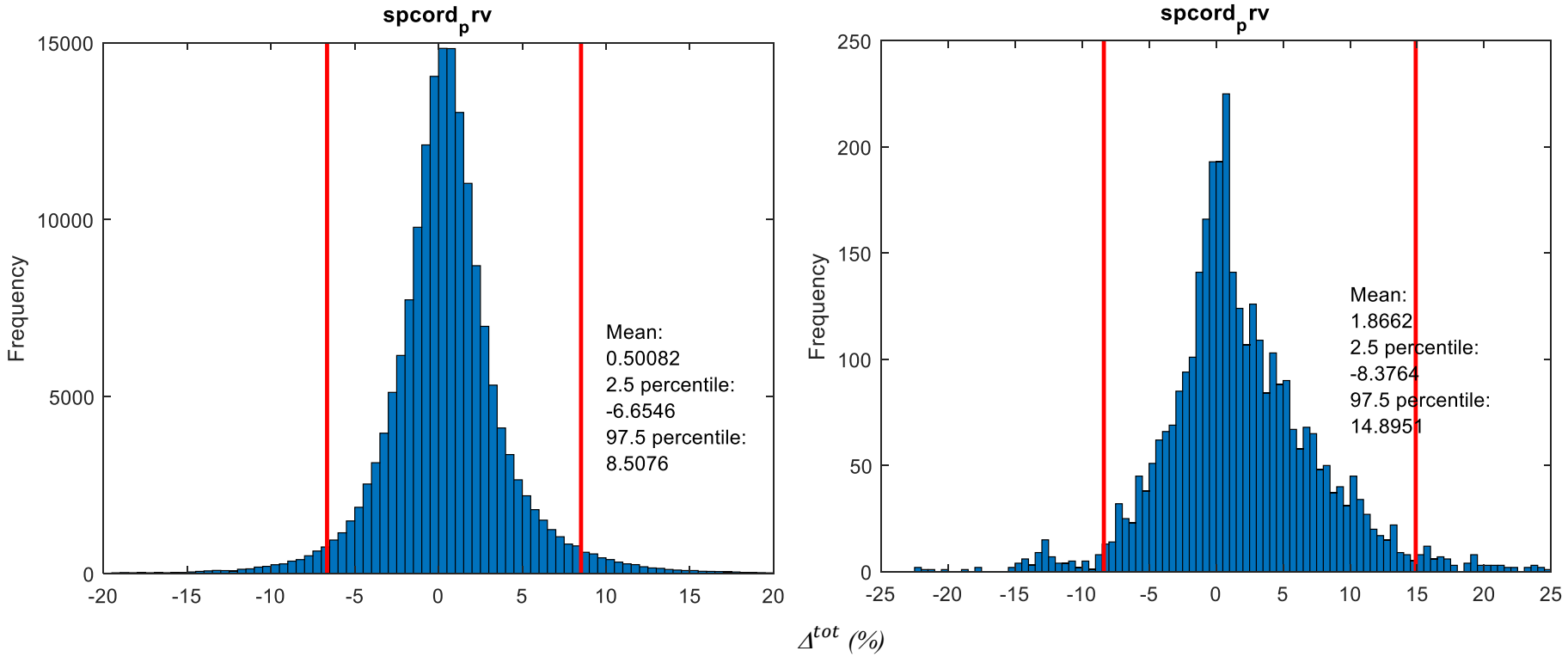


Figure A.13.8: Histogram showing the distribution of Δ^{tot} for the spinal cord planning risk volume (PRV) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

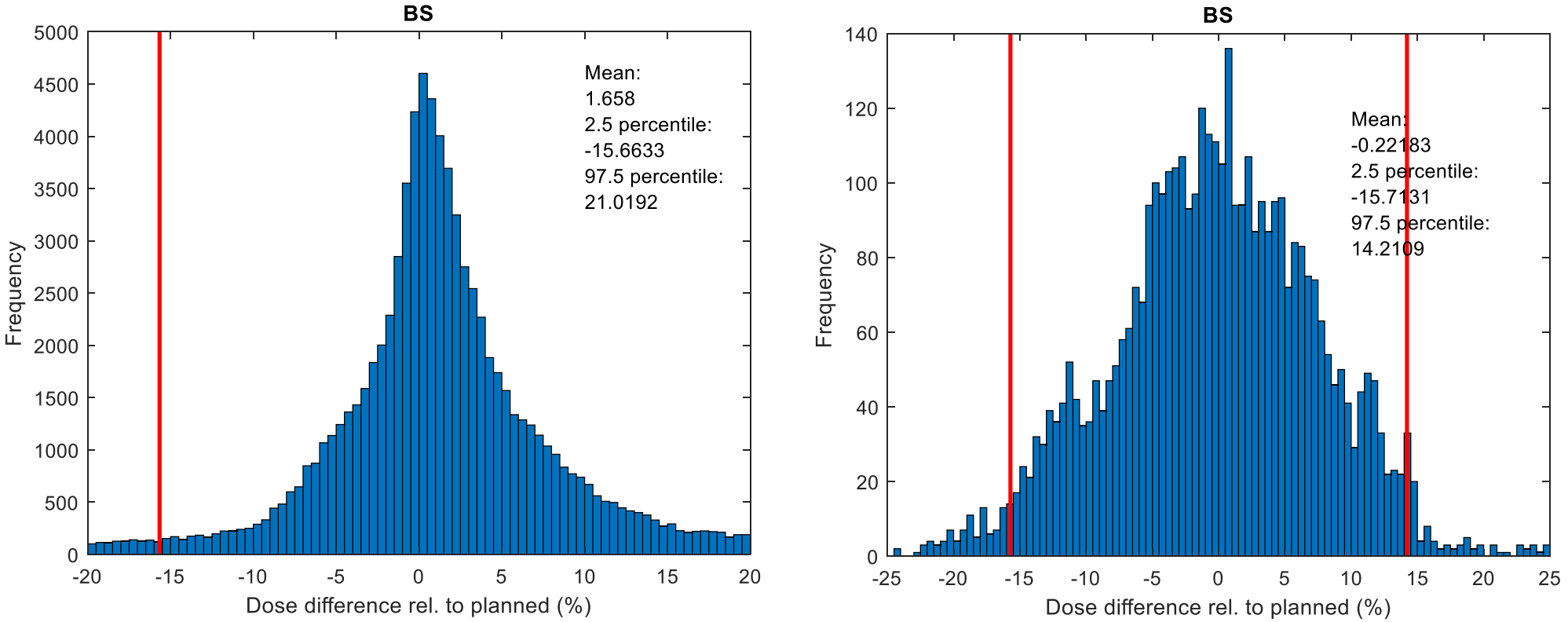


Figure A.13.9: Histogram showing the distribution of Δ^{tot} for the brainstem of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

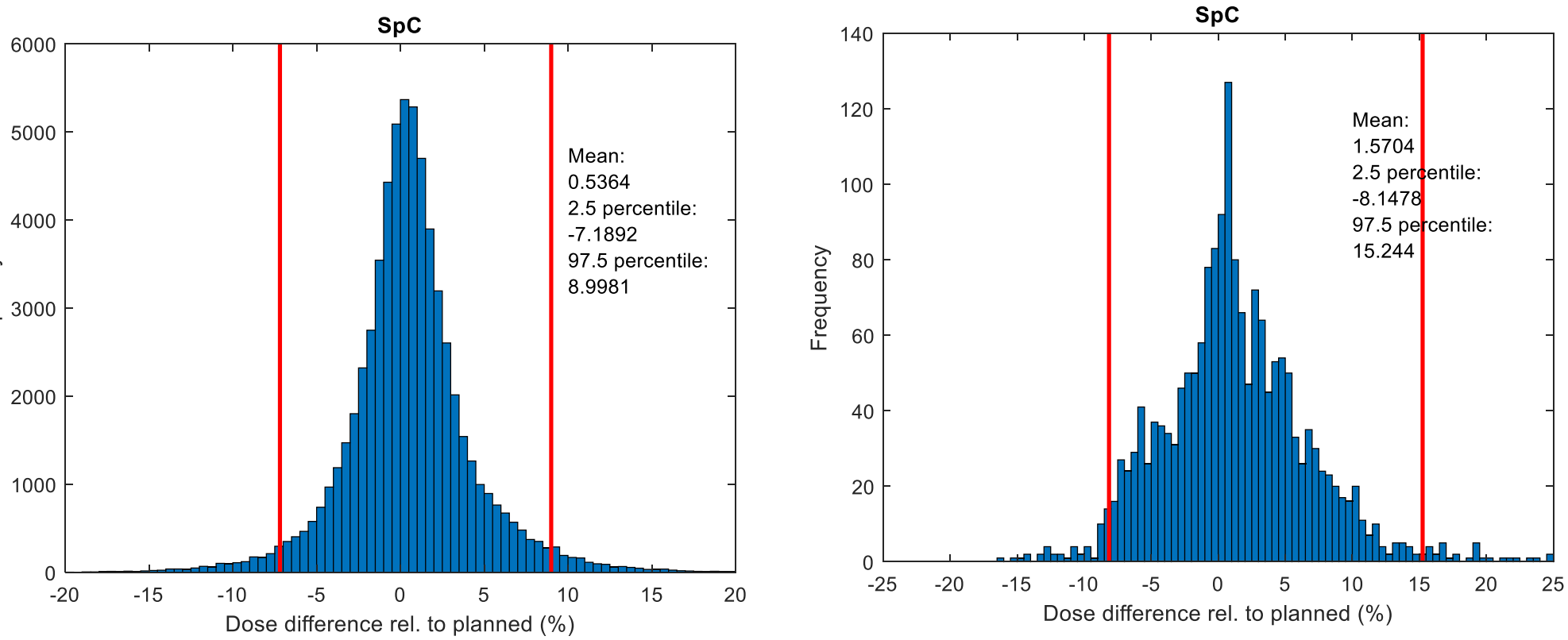


Figure A.13.10: Histogram showing the distribution of Δ^{tot} for the spinal cord of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

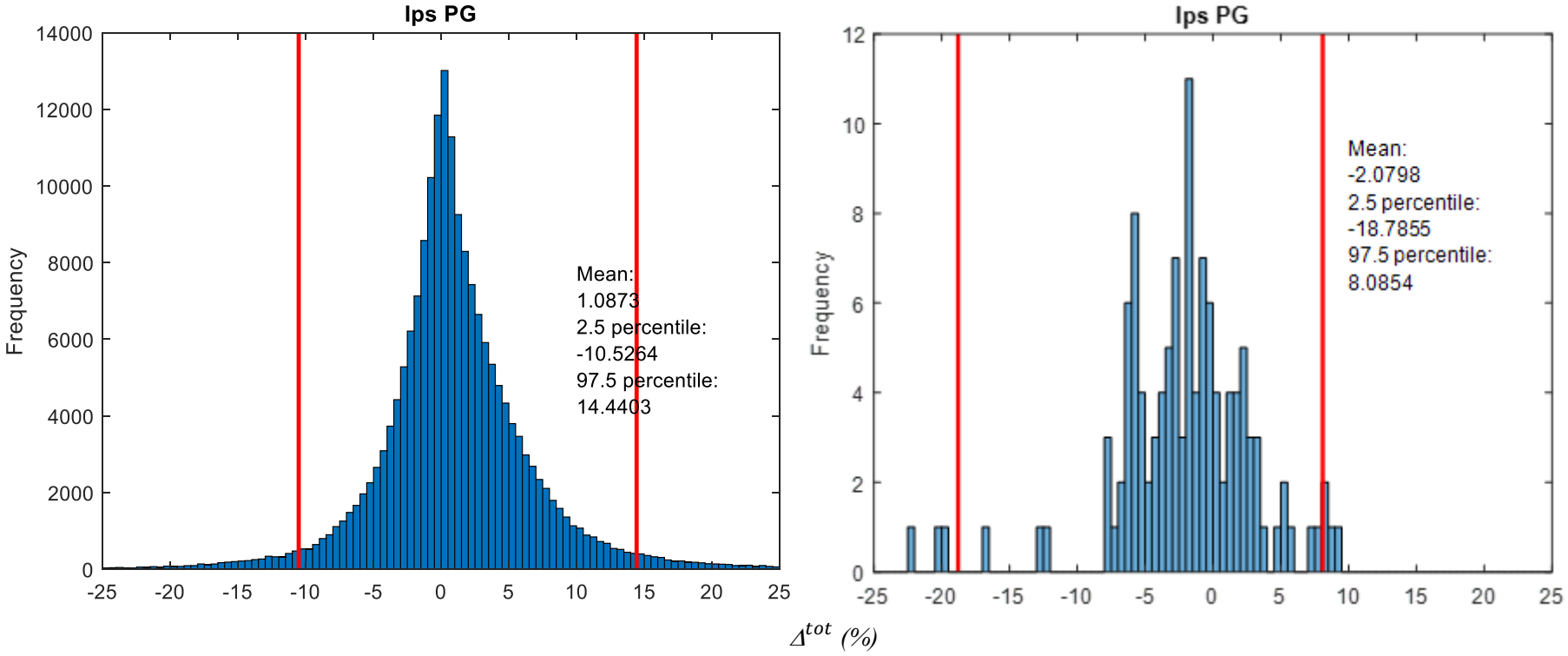


Figure A.13.11: Histogram showing the distribution of Δ^{tot} for the ipsilateral parotid gland (PG) of all patients for inverse consistent (left) and non- inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

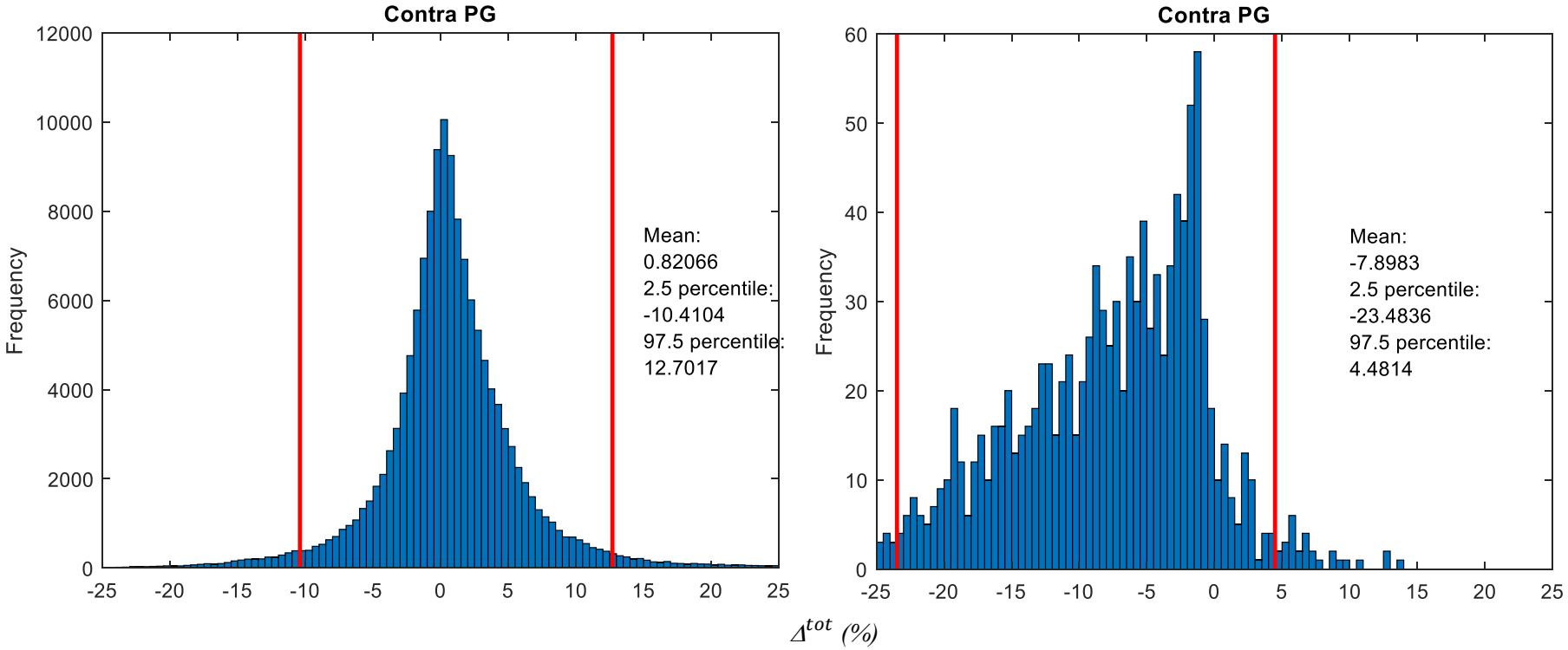


Figure A.13.12: Histogram showing the distribution of Δ^{tot} for the contralateral parotid gland (PG) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

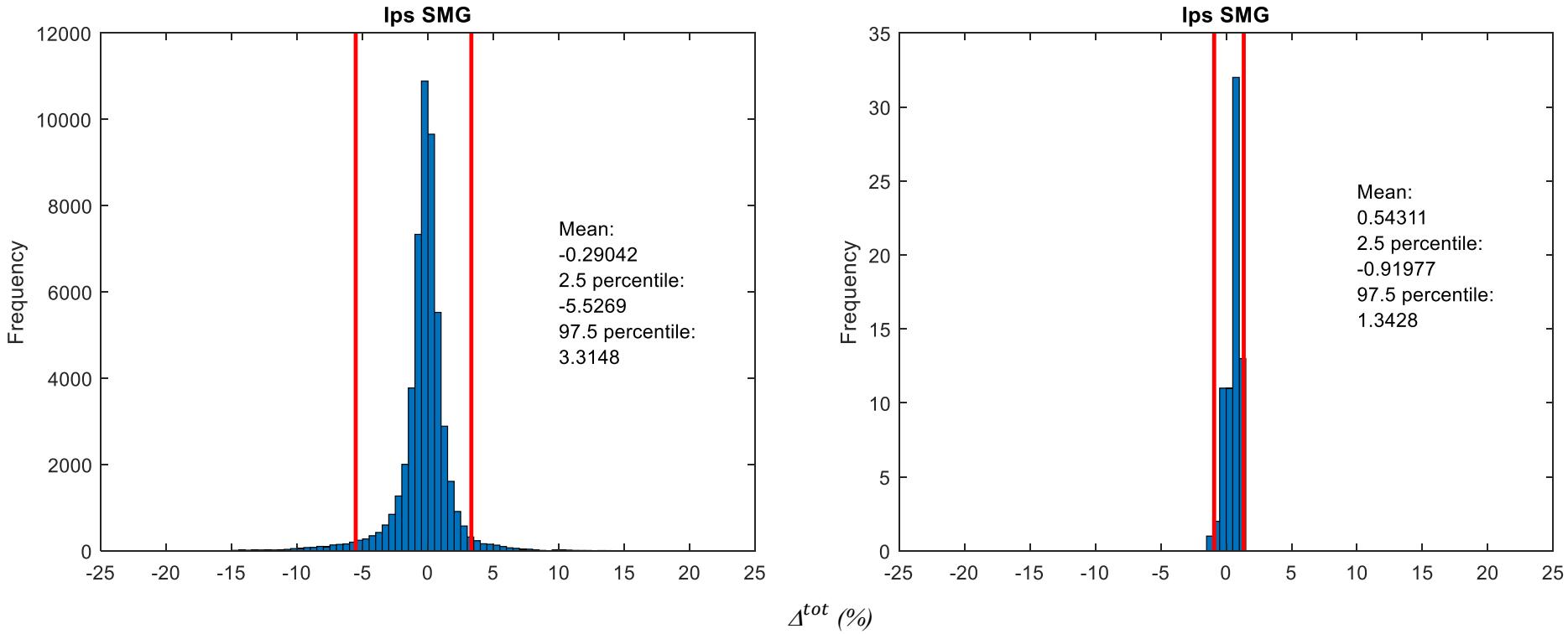


Figure A.13.13: Histogram showing the distribution of Δ^{tot} for the ipsilateral submandibular gland (SMG) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

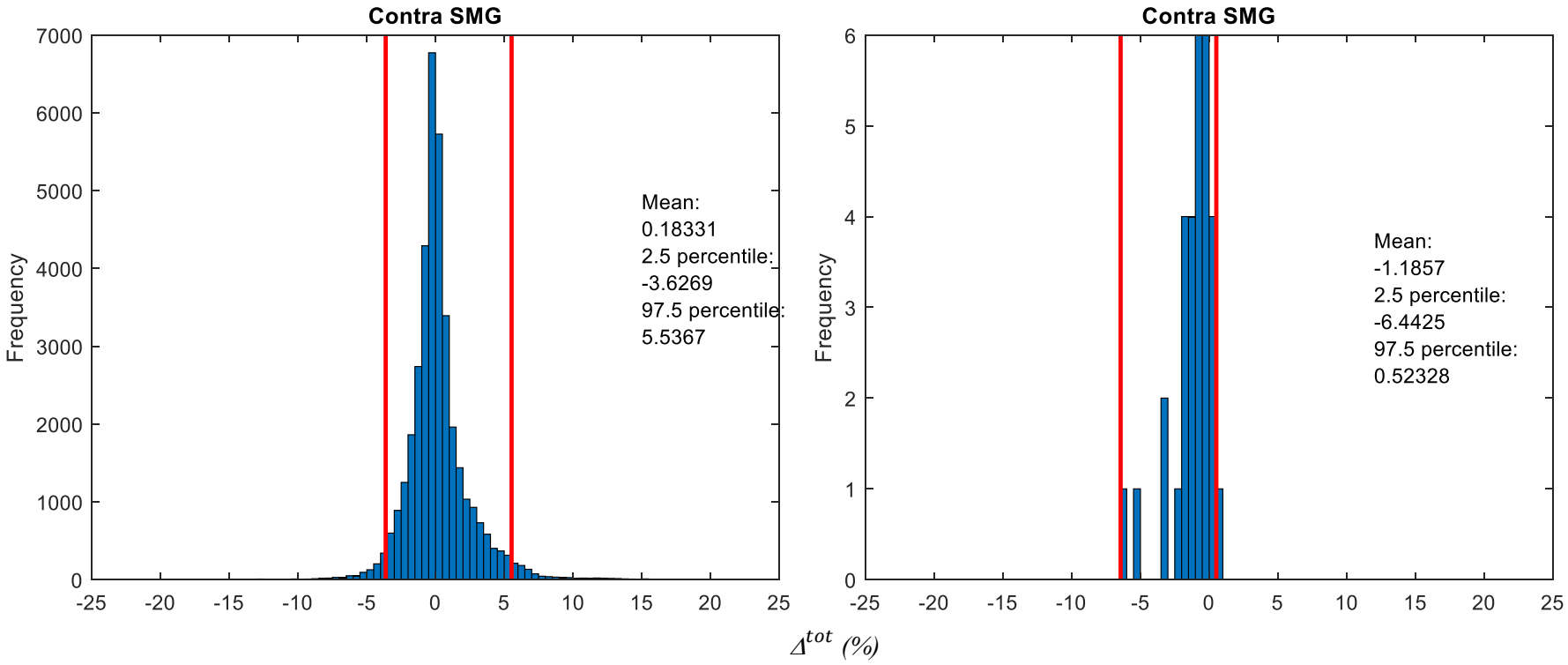


Figure A.13.14: Histogram showing the distribution of Δ^{tot} for the contralateral submandibular gland (SMG) of all patients for inverse consistent (left) and inverse inconsistent (right) voxels. Red vertical lines indicate the 95% percentile range.

A.14: Structures available for analysis

Table A.14: Patient-specific structure count and structure-specific dose-volume histogram (DVH) analysis parameter. CTV = clinical target volume. PTV = planning target volume. PRV = planning risk volume. PG = parotid gland. SMG = submandibular gland. $D_{99\%}$ = minimum dose to 99% of volume. $D_{98\%}$ = minimum dose to 98% of volume. $D_{0.1cc}$ = maximum dose to 0.1 cubic centimeters. D_{mean}= mean dose.

	Patient	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***n***
Structure	***DVH parameter***													
CTV 54	$D_{99\%}$	X	X	X	X	X	X	X	X	X	X	X	X	**12**
CTV 60	$D_{99\%}$						X			X	X		X	**4**
CTV 66	$D_{99\%}$	X		X	X	X	X	X	X	X	X	X	X	**11**
PTV 54	$D_{98\%}$	X	X	X	X	X	X	X	X	X	X	X	X	**12**
PTV 60	$D_{98\%}$						X			X	X		X	**4**
PTV 66	$D_{98\%}$	X		X	X	X	X	X	X	X	X	X	X	**11**
Spinal cord	$D_{0.1cc}$	X	X	X	X	X	X	X	X	X	X	X	X	**12**
Brainstem	$D_{0.1cc}$	X	X	X	X	X	X	X	X	X	X	X	X	**12**
Spinal cord PRV	$D_{0.1cc}$	X	X	X	X	X	X	X	X	X	X	X	X	**12**
Brainstem PRV	$D_{0.1cc}$	X	X	X	X	X	X	X	X	X	X	X	X	**12**
Ipsilateral PG	D_{mean}	X	X	X	X	X	X	X	X	X	X	X	X	**14**
								X		X				
Contralateral PG	D_{mean}	X	X	X	X	X	X		X		X	X	X	**10**
Ipsilateral SMG	D_{mean}	X	X	X	X	X	X	X	X	X	X		X	**13**
								X		X				
Contralateral SMG	D_{mean}	X	X	X	X	X	X		X		X	X		**9**
Total														**148**

A.15: Generation of local volumes

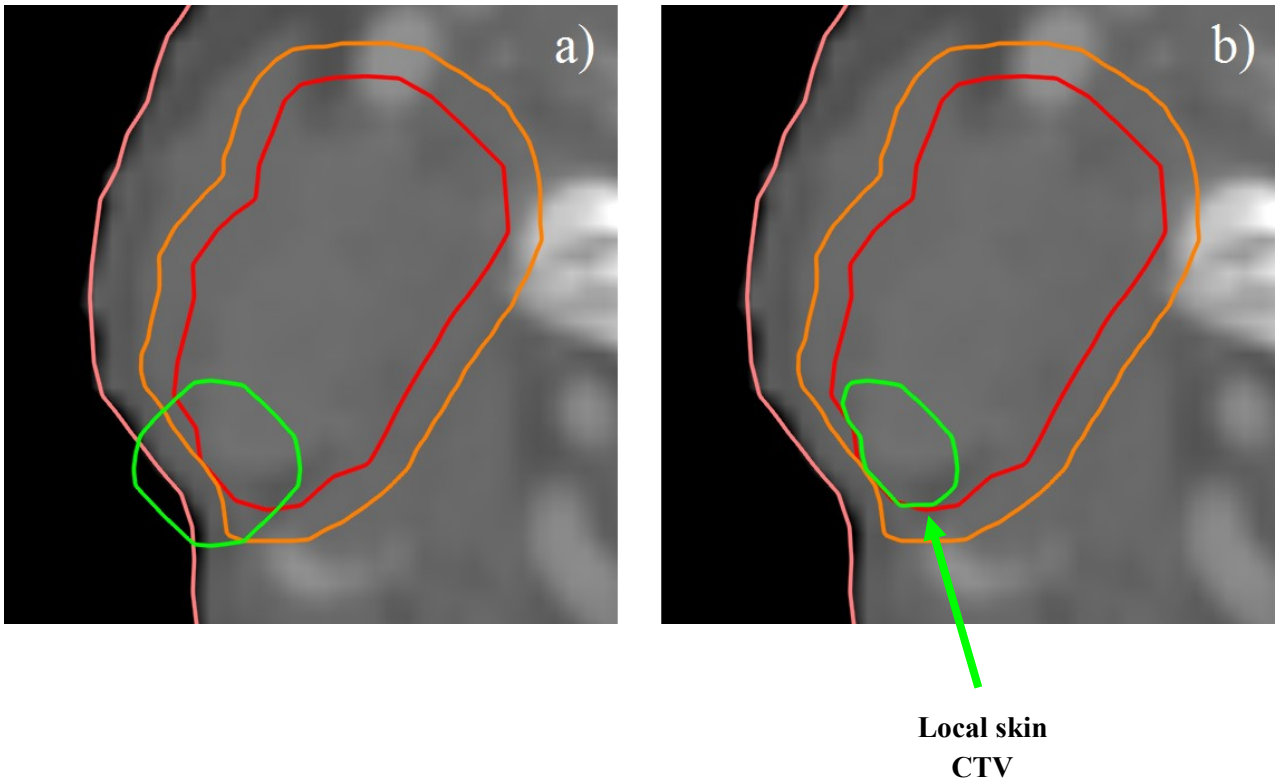


Figure A.15.1: Generation of the local skin clinical target volume (CTV). (a) A 1 cm diameter sphere (green contour) was generated at the point of minimum distance between the CTV (red contour) and skin (pink contour). (b) The sub-volume of the sphere that overlapped with the CTV was defined as the local skin CTV. The 3 mm planning target volume (PTV) margin is shown by the orange contour. Note that the PTV and CTV are both cropped 3 mm back from the skin due to dose calculation unreliability in this region.

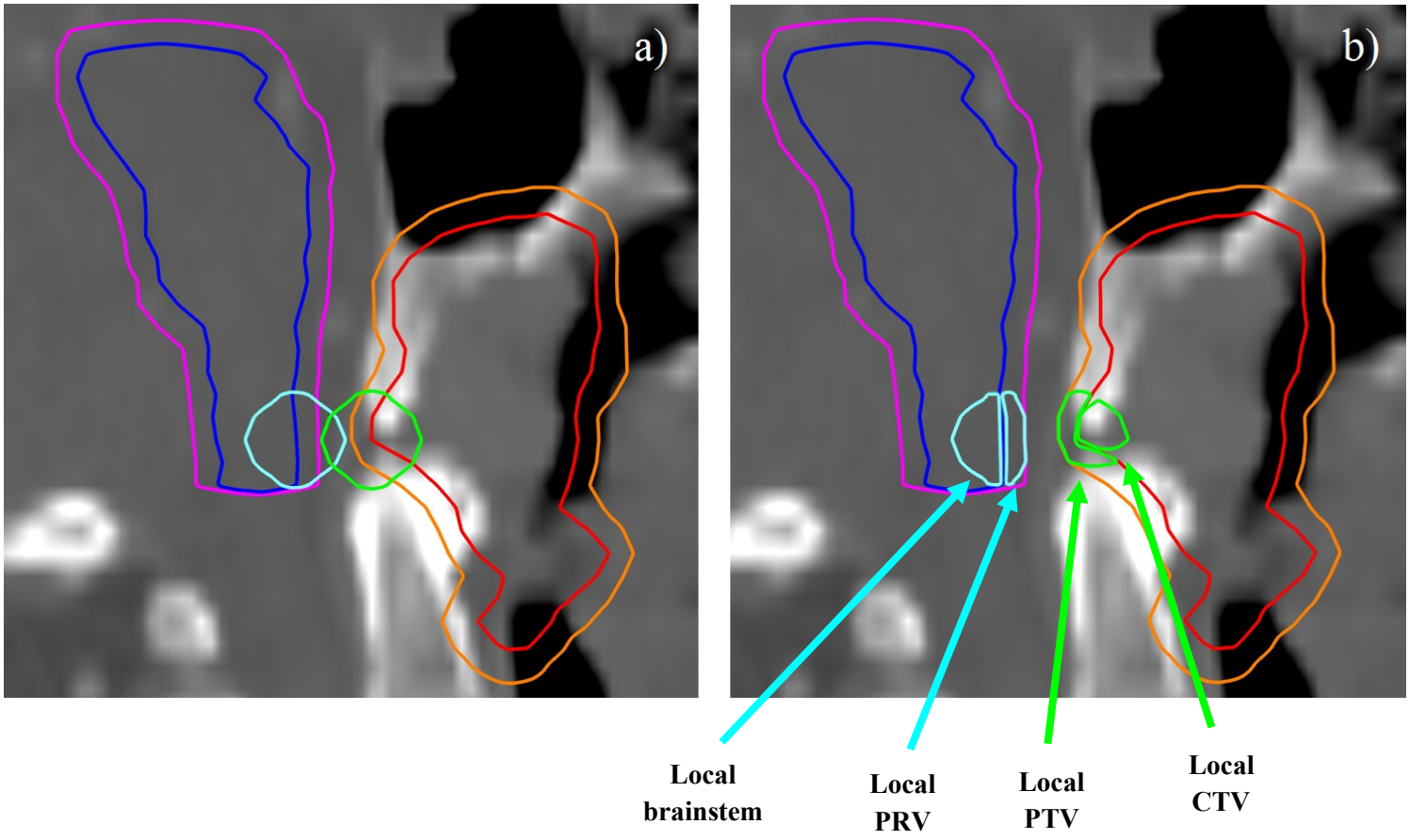


Figure A.15.2: Generation of the local critical organ at risk (OAR) volumes. (a) Two 1 cm diameter spheres (green contour and teal contour) were generated. One sphere was centred on the high-dose clinical target volume (CTV) (red contour) boundary and one sphere on the brainstem (blue contour) boundary at the point of minimum distance between the CTV and brainstem. (b) The sub-volumes of the spheres that overlapped with the brainstem and CTV were defined as the local brainstem and local CTV, respectively. The sub-volumes of the spheres that overlapped with the planning risk volume (PRV) margin (magenta contour) and planning target volume (PTV) margin (orange contour) were defined as the local PRV and local PTV, respectively. The generation of the local volumes was repeated for 5 mm PRV/PTV margin plans and was also repeated for the spinal cord.

A.16: Normality tests

The planned, accumulated and *ΔD* DVH parameters were tested for normality for the LD-CTVs (i.e., CTV 54), HD-CTVs (i.e., CTV 66) and the ipsilateral PGs. It was determined that the planned, accumulated and *ΔD* DVH parameters were not always normally distributed and further normality tests of these parameters for the remaining structures was not conducted. However, the *differences* between the DVH parameters of the 5 and 3 mm plans were normally distributed. Supplementary Material E.1 presents the normality results for the planned, accumulated and *ΔD* DVH parameters for the LD-CTV, HD-CTV and the ipsilateral PG. Supplementary Material E.2 presents the results of the normality tests for the *differences* between the DVH parameters of the 5 and 3 mm plans. Q-Q plots and Shapiro-Wilk tests were used to test for normality and are presented below.

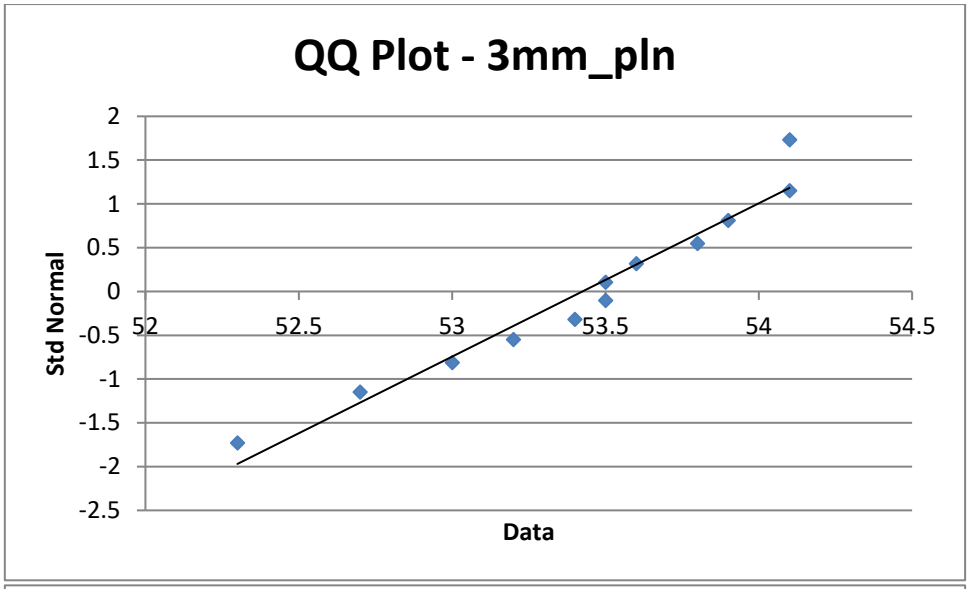
Supplementary Material E.1

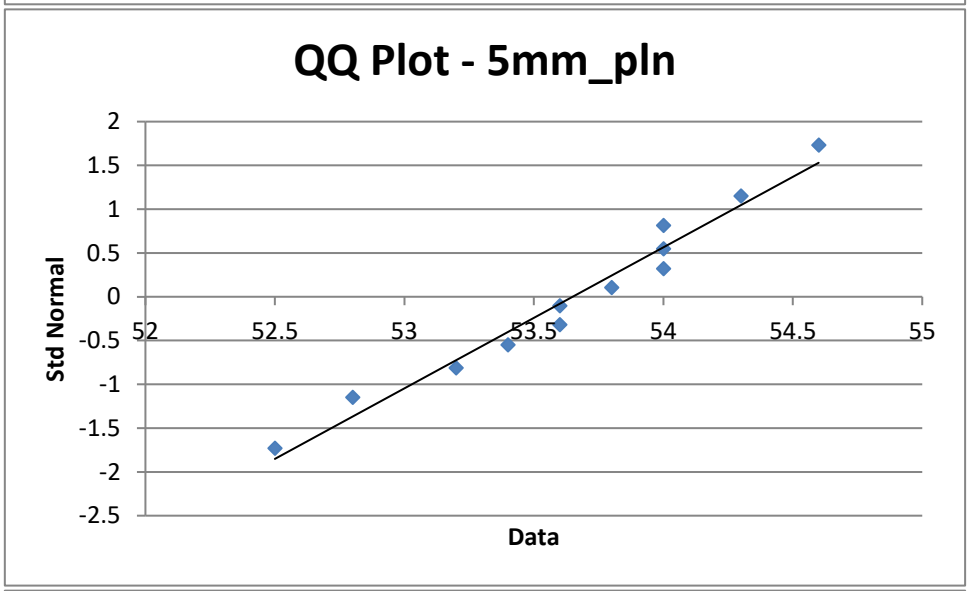
3mm_pln = Treatment plan with 3 mm PTV margins; the planned dose
3mm_acc = Treatment plan with 3 mm PTV margins; the accumulated dose
3mm_dD = Treatment plan with 3 mm PTV margins; the change in dose (i.e., acc – pln)

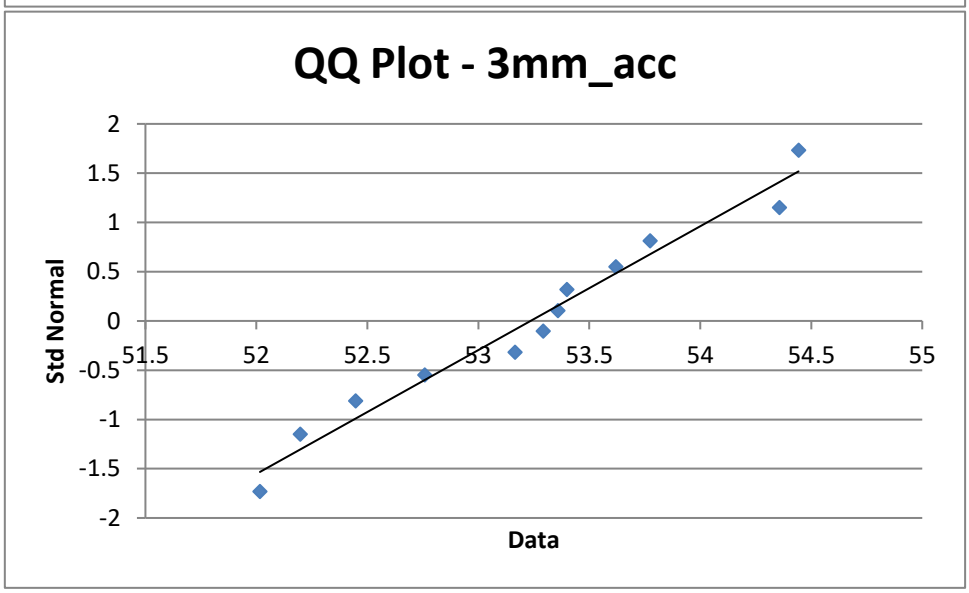
5mm_pln = Treatment plan with 5 mm PTV margins; the planned dose
5mm_acc = Treatment plan with 5 mm PTV margins; the accumulated dose
5mm_dD = Treatment plan with 5 mm PTV margins; the change in dose (i.e., acc – pln)

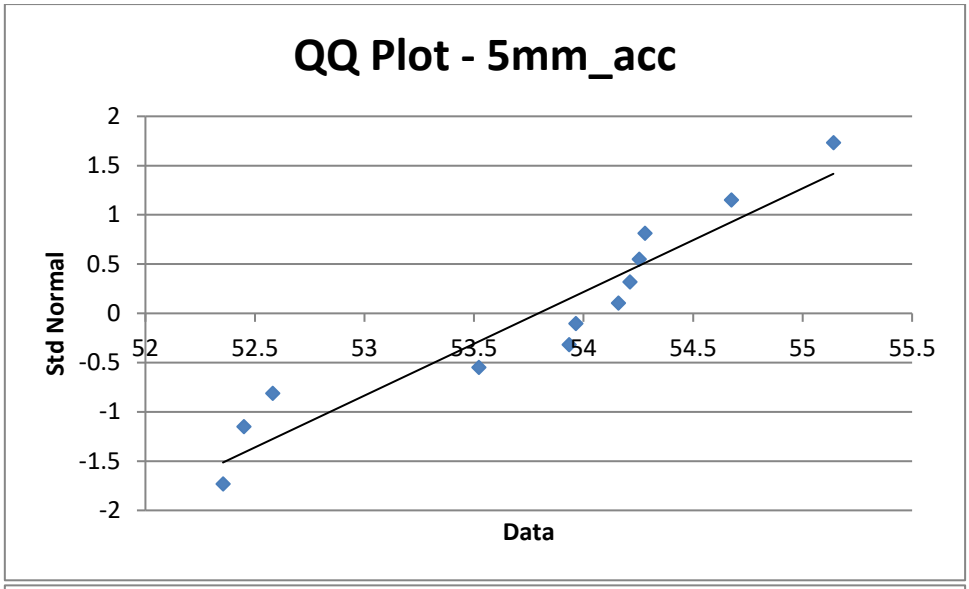
CTV 54

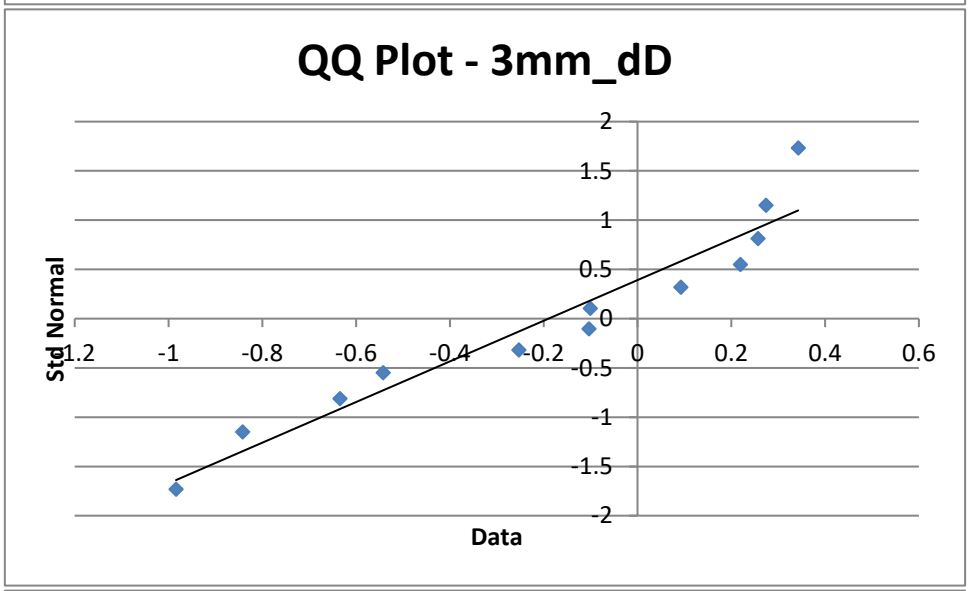
Shapiro-Wilk Test						
	3mm_pln	*5mm_pln*	*3mm_acc*	*5mm_acc*	*3mm_dD*	*5mm_dD*
W-stat	0.943525	0.967372	0.95848	0.896318	0.909193	0.848788
p-value	0.545042	0.881424	0.761936	0.142176	0.208325	0.035471
alpha	0.05	0.05	0.05	0.05	0.05	0.05
normal	yes	yes	yes	yes	yes	no

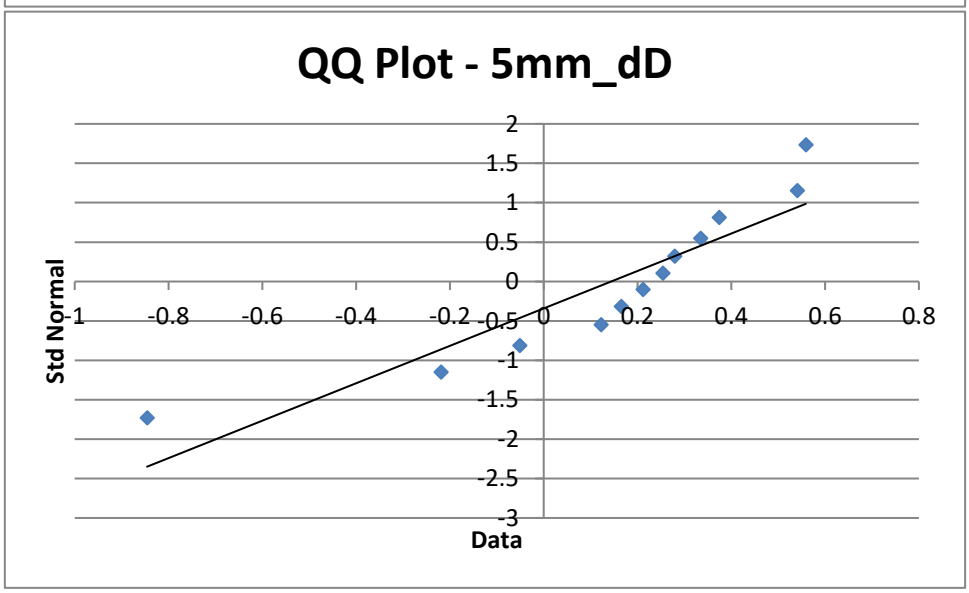

QQ Plot - 3mm_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
52
52.5
53
53.5
54
54.5


QQ Plot - 5mm_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
52
52.5
53
53.5
54
54.5
55


QQ Plot - 3mm_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
51.5
52
52.5
53
53.5
54
54.5
55

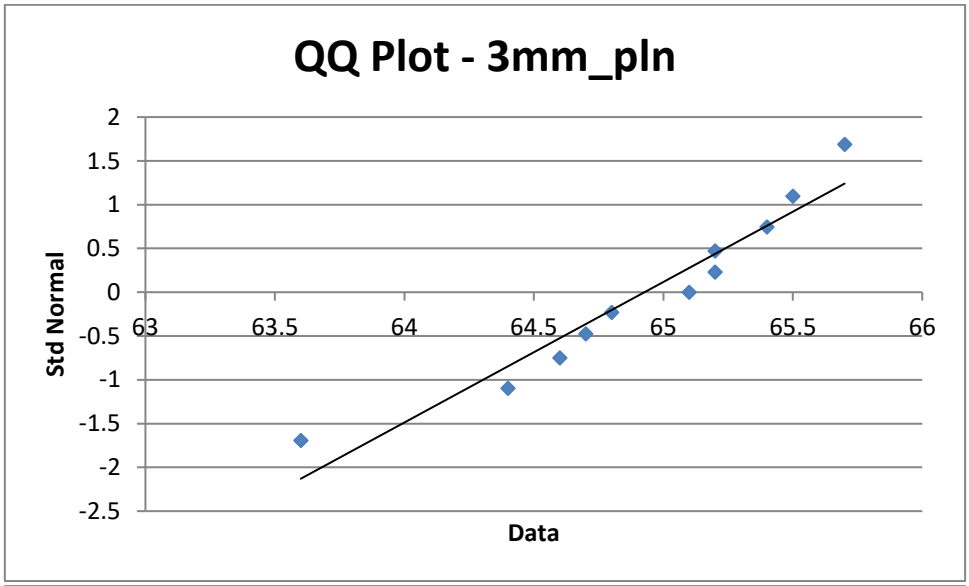

QQ Plot - 5mm_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
52
52.5
53
53.5
54
54.5
55
55.5

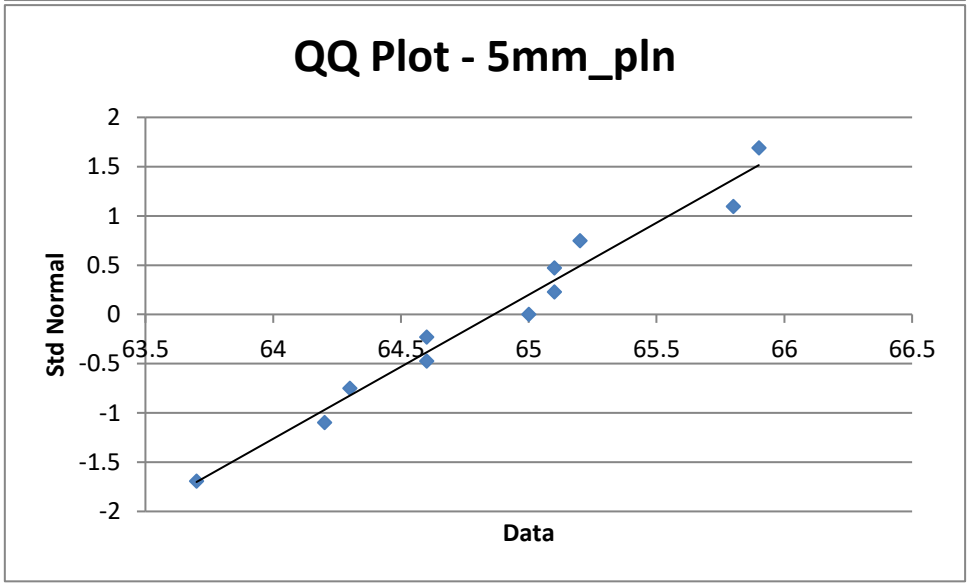

QQ Plot - 3mm_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-1.2
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6

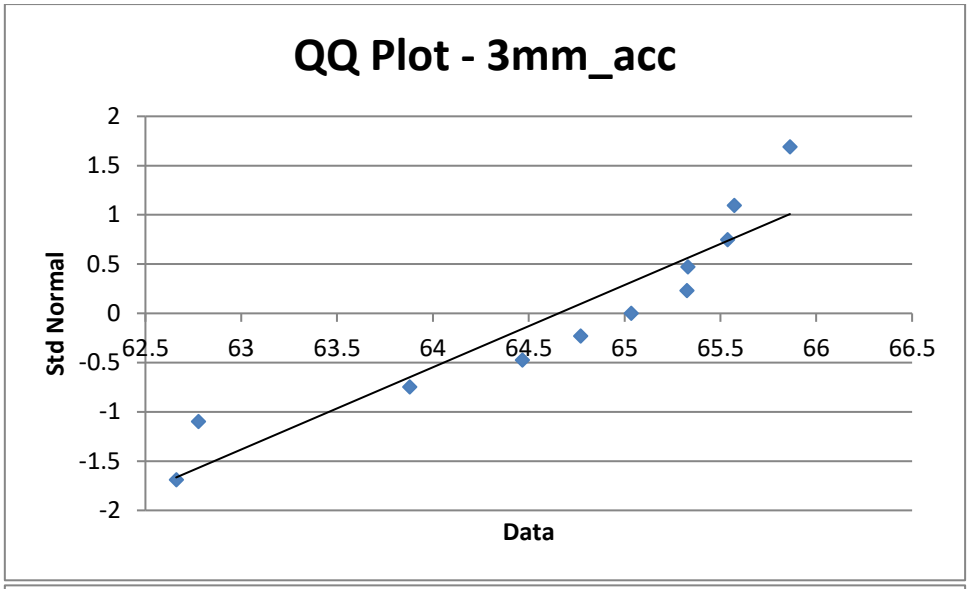

QQ Plot - 5mm_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8

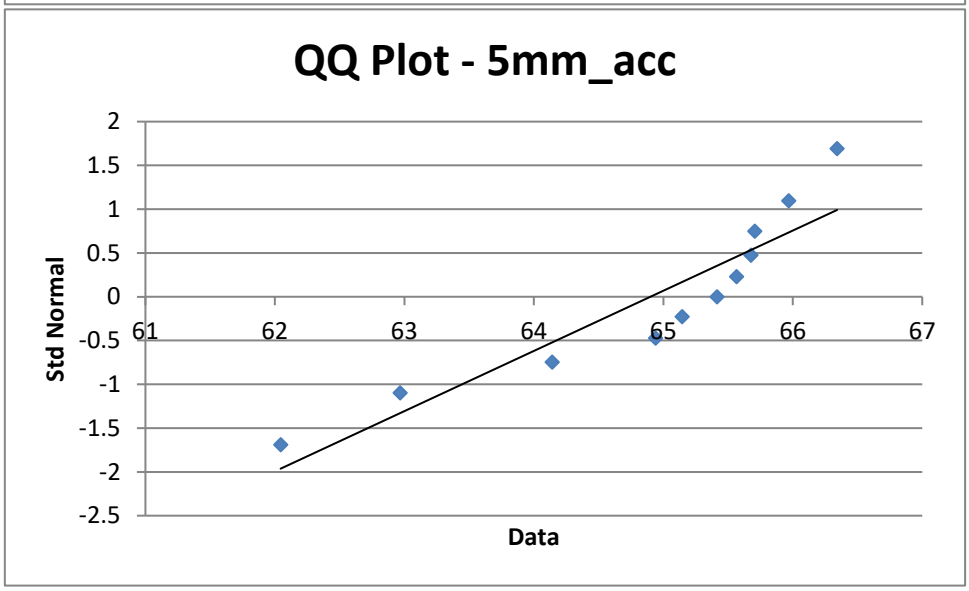
CTV 54

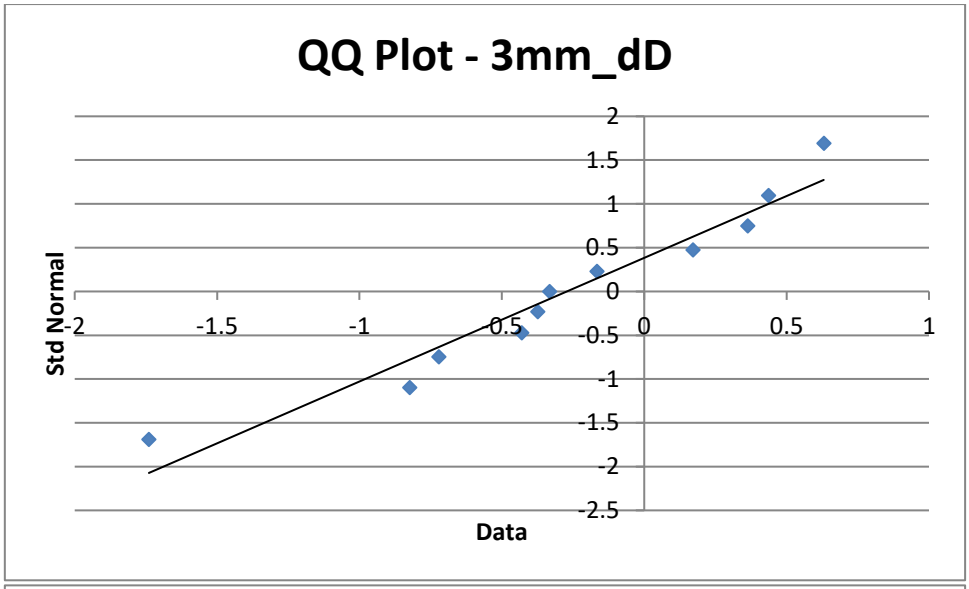
Shapiro-Wilk Test						
	3mm_pln	*5mm_pln*	*3mm_acc*	*5mm_acc*	*3mm_dD*	*5mm_dD*
W-stat	0.936317	0.966111	0.863591	0.854681	0.944933	0.727134
p-value	0.478216	0.844804	0.06409	0.049112	0.580096	0.001048
alpha	0.05	0.05	0.05	0.05	0.05	0.05
normal	yes	yes	yes	no	yes	no

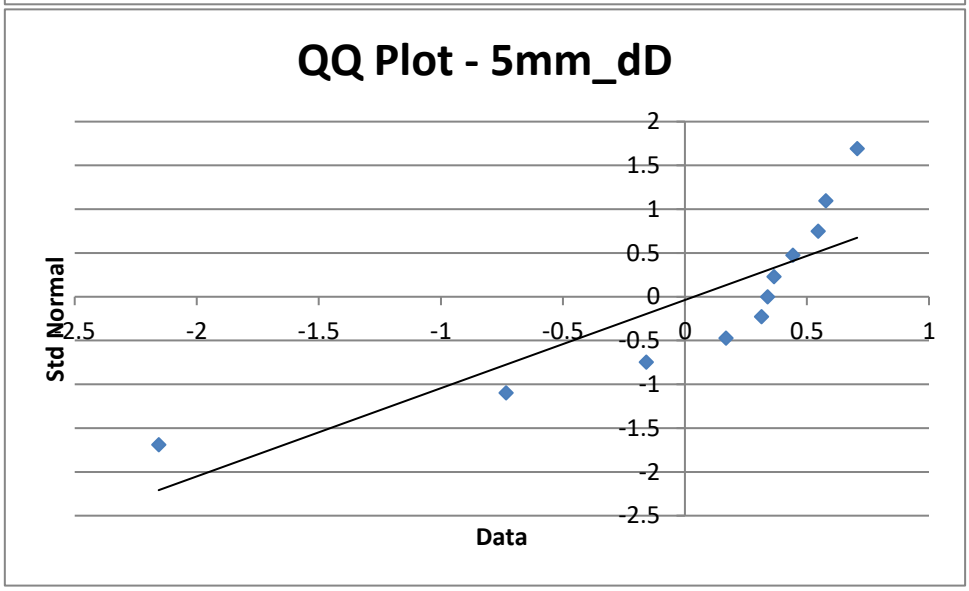





QQ Plot - 3mm_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
62.5
63
63.5
64
64.5
65
65.5
66
66.5

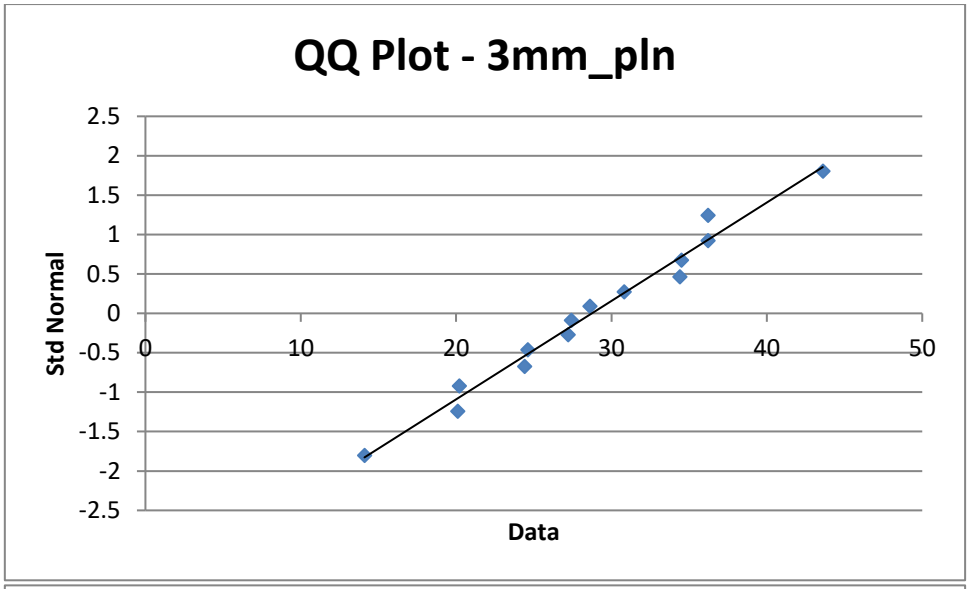

QQ Plot - 5mm_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
61
62
63
64
65
66
67

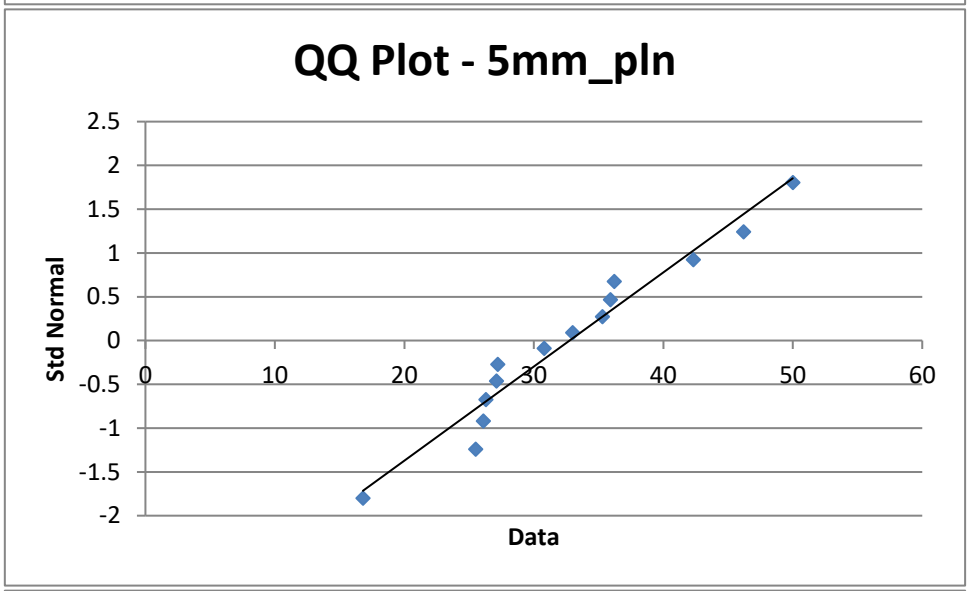


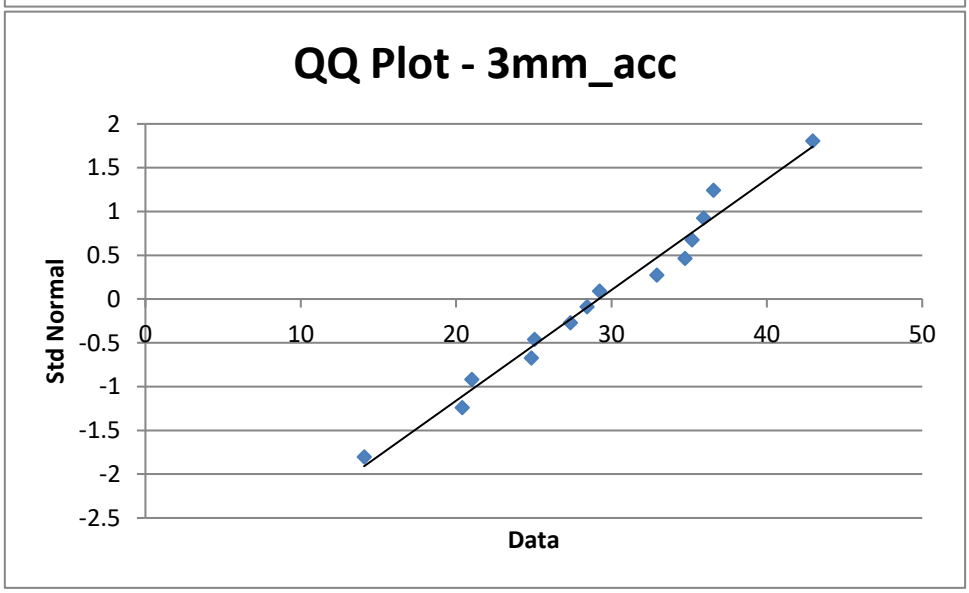


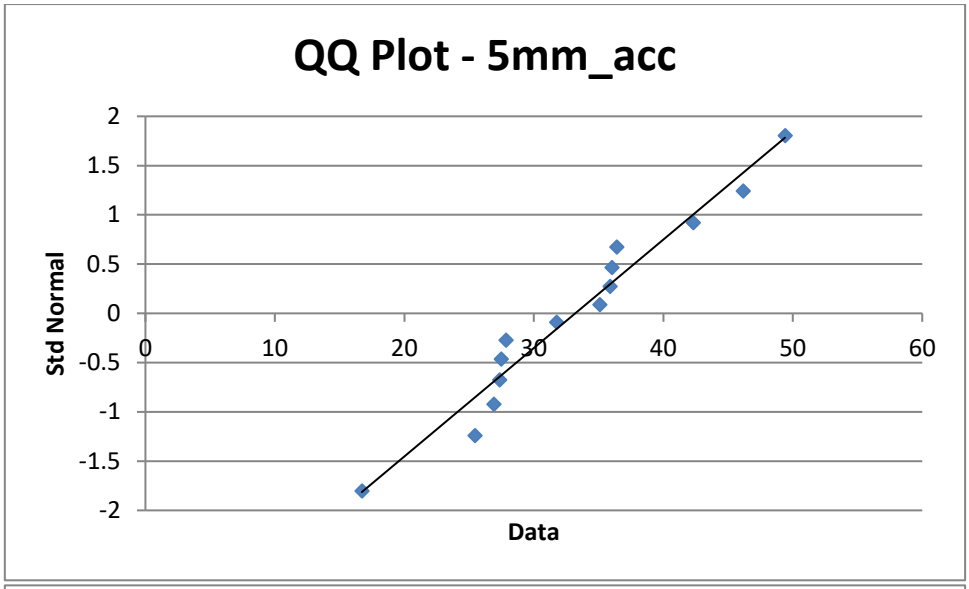
Ips PG

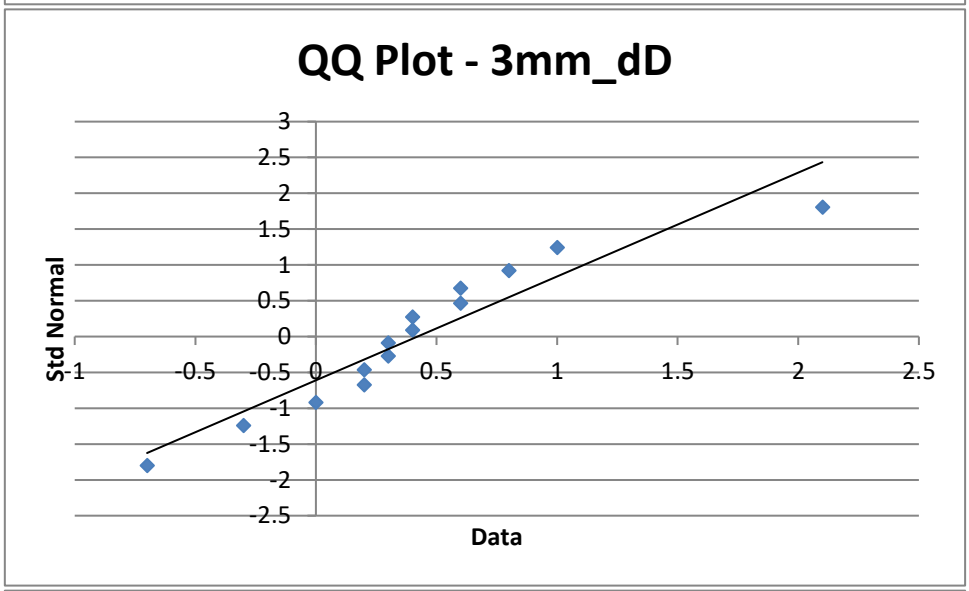
Shapiro-Wilk Test						
	3mm_pln	*5mm_pln*	*3mm_acc*	*5mm_acc*	*3mm_dD*	*5mm_dD*
W-stat	0.98194	0.956513	0.979519	0.960808	0.913207	0.926384
p-value	0.984517	0.665489	0.971924	0.736383	0.175411	0.271321
alpha	0.05	0.05	0.05	0.05	0.05	0.05
normal	yes	yes	yes	yes	yes	yes

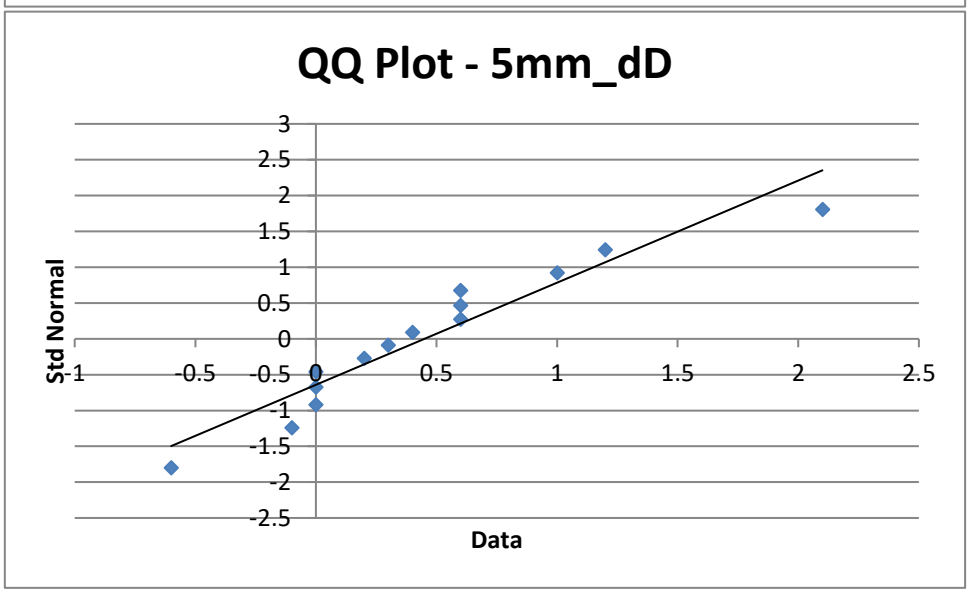

QQ Plot - 3mm_pln
Std Normal
Data
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
0
10
20
30
40
50


QQ Plot - 5mm_pln
Std Normal
Data
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
10
20
30
40
50
60


QQ Plot - 3mm_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
0
10
20
30
40
50


QQ Plot - 5mm_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
10
20
30
40
50
60


QQ Plot - 3mm_dD
Std Normal
Data
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-1
-0.5
0
0.5
1
1.5
2
2.5


QQ Plot - 5mm_dD
Std Normal
Data
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-1
-0.5
0
0.5
1
1.5
2
2.5

Supplementary Material E.2

Shapiro-Wilk Test												
	CTV54_pln	*CTV60_pln*	*CTV66_pln*	*PTV54_pln*	*PTV60_pln*	*PTV66_pln*	*SpCPRV_pln*	*BSPRV_pln*	*Ips PG_pln*	*Con PG_pln*	*Ips SMG_pln*	*Con SMG_pln*
W-stat	*0.9573515*	*0.7172207*	*0.97271705*	*0.9093418*	*0.8282964*	*0.9358162*	*0.935636113*	*0.89209911*	*0.96779665*	*0.94315548*	*0.922820162*	*0.85795049*
p-value	0.7454602	0.0180998	0.91255958	0.2092412	0.1634032	0.4726295	0.443686319	0.12541955	0.84587468	0.588629064	0.273753477	0.091079879
alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
normal	yes	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

CTV_54_pln = Differences between the $D_{99\%}$ of the 5 and 3 mm *planned* doses for the LD-CTV

CTV_60_pln = Differences between the $D_{99\%}$ of the 5 and 3 mm *planned* doses for the ID-CTV

CTV_66_pln = Differences between the $D_{99\%}$ of the 5 and 3 mm *planned* doses for the HD-CTV

PTV_54_pln = Differences between the $D_{98\%}$ of the 5 and 3 mm *planned* doses for the LD-PTV

PTV_60_pln = Differences between the $D_{98\%}$ of the 5 and 3 mm *planned* doses for the ID-PTV

PTV_66_pln = Differences between the $D_{98\%}$ of the 5 and 3 mm *planned* doses for the HD-PTV

SpCPRV_pln = Differences between the $D_{0.1cc}$ of the 5 and 3 mm *planned* doses for the spinal cord PRV

BSPRV_pln = Differences between the $D_{0.1cc}$ of the 5 and 3 mm *planned* doses for the brainstem PRV

Ips PG_pln = Differences between the D_{mean} of the 5 and 3 mm *planned* doses for the ipsilateral parotid gland

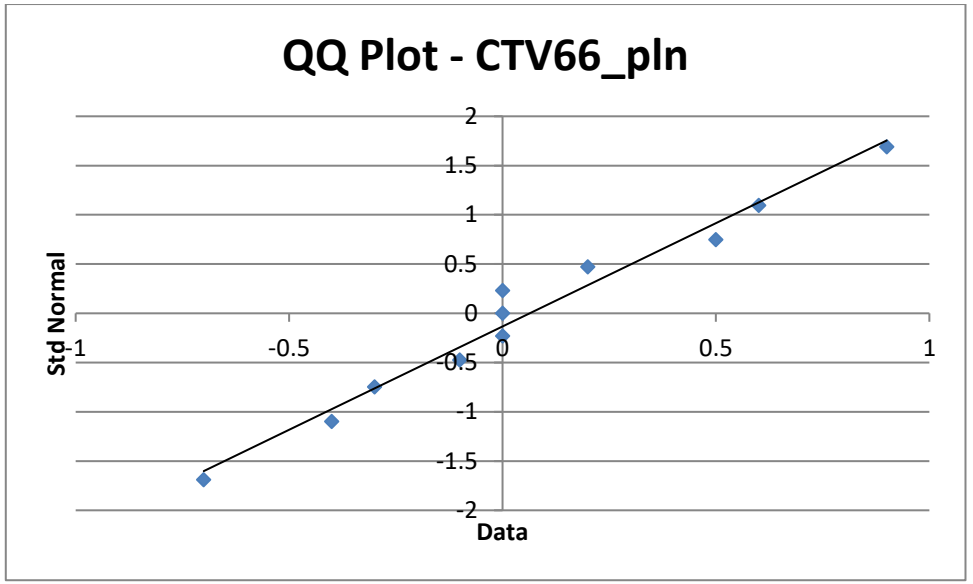
Con PG_pln = Differences between the D_{mean} of the 5 and 3 mm *planned* doses for the contralateral parotid gland

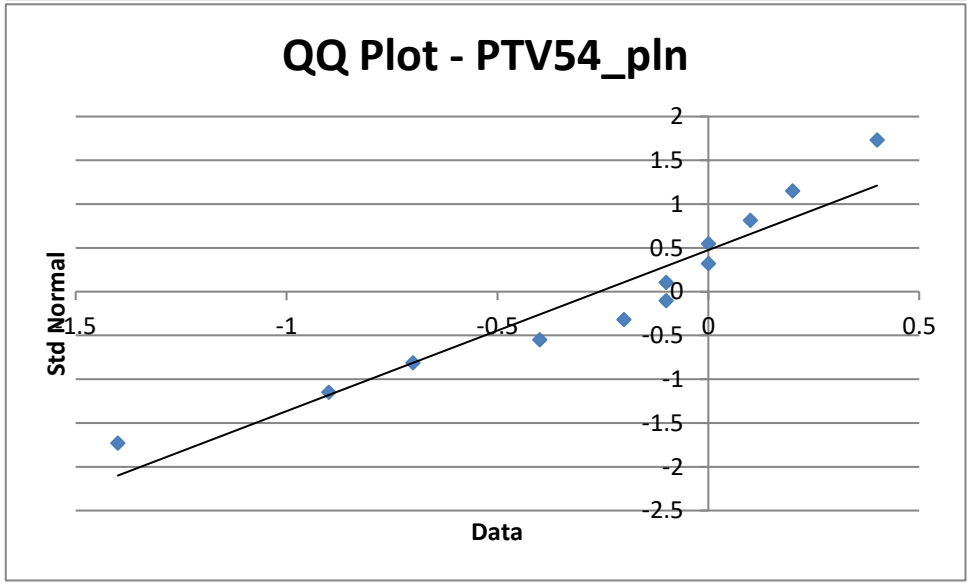
Ips SMG_pln = Differences between the D_{mean} of the 5 and 3 mm *planned* doses for the ipsilateral submandibular gland

Con SMG_pln = Differences between the D_{mean} of the 5 and 3 mm *planned* doses for the contralateral submandibular gland

QQ Plot - CTV54_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-1.5
-1
-0.5
0
0.5
1

QQ Plot - CTV60_pln
Std Normal
Data
1.5
1
0.5
0
-0.5
-1
-1.5
-1.5
-1
-0.5
0
0.5


QQ Plot - CTV66_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-1
-0.5
0
0.5
1


QQ Plot - PTV54_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-1.5
-1
-0.5
0
0.5

QQ Plot - PTV60_pln
Std Normal
Data
1.5
1
0.5
0
-0.5
-1
-1.5
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5
0.6

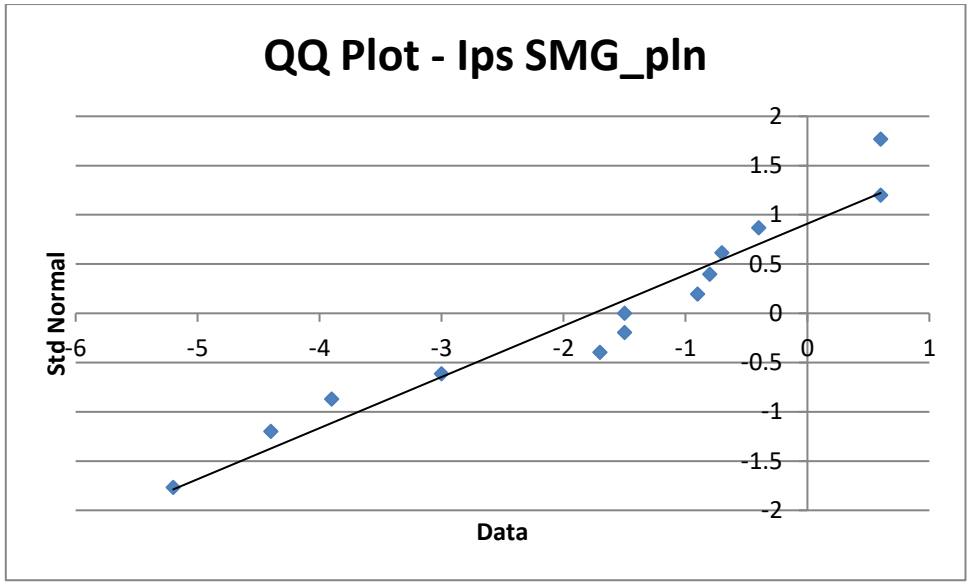
QQ Plot - PTV66_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

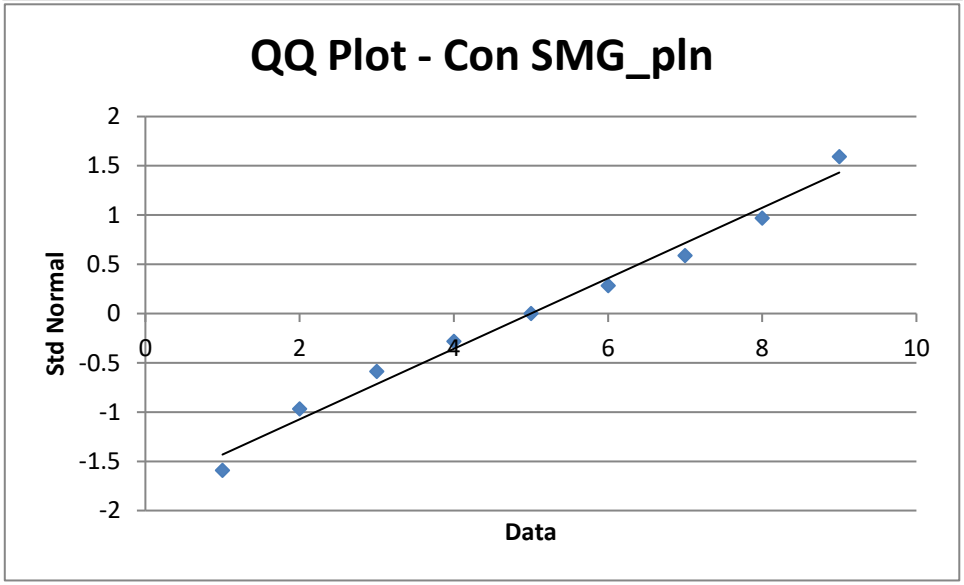
QQ Plot - SpCPRV_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-12
-10
-8
-6
-4
-2
0
2
4

QQ Plot - BSPRV_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-25
-20
-15
-10
-5
0
5

QQ Plot - Ips PG_pln
Std Normal
Data
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-12
-10
-8
-6
-4
-2
0
2
4

QQ Plot - Con PG_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-6
-5
-4
-3
-2
-1
0


QQ Plot - Ips SMG_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-6
-5
-4
-3
-2
-1
0
1


QQ Plot - Con SMG_pln
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
2
4
6
8
10

Shapiro-Wilk Test												
	CTV54_acc	*CTV60_acc*	*CTV66_acc*	*PTV54_acc*	*PTV60_acc*	*PTV66_acc*	*SpCPRV_acc*	*BSPRV_acc*	*Ips PG_acc*	*Con PG_acc*	*Ips SMG_acc*	*Con SMG_acc*
W-stat	*0.9431502*	*0.94510135*	*0.96321135*	*0.9128057*	*0.86540617*	*0.894349*	*0.93870526*	*0.8046968*	*0.9728651*	*0.93631702*	*0.92437968*	*0.82682767*
p-value	*0.5*	*0.7*	*0.8*	*0.2*	*0.3*	*0.2*	*0.5*	*0.0*	*0.9*	*0.5*	*0.3*	*0.0*
alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
normal	yes	yes	yes	yes	yes	yes	yes	no	yes	yes	yes	no

CTV_54_acc = Differences between the $D_{99\%}$ of the 5 and 3 mm *accumulated* doses for the LD-CTV

CTV_60_acc = Differences between the $D_{99\%}$ of the 5 and 3 mm *accumulated* doses for the ID-CTV

CTV_66_acc = Differences between the $D_{99\%}$ of the 5 and 3 mm *accumulated* doses for the HD-CTV

PTV_54_ acc = Differences between the $D_{98\%}$ of the 5 and 3 mm *accumulated* doses for the LD-PTV

PTV_60_acc = Differences between the $D_{98\%}$ of the 5 and 3 mm *accumulated* doses for the ID-PTV

PTV_66_acc = Differences between the $D_{98\%}$ of the 5 and 3 mm *accumulated* doses for the HD-PTV

SpCPRV_acc = Differences between the $D_{0.1cc}$ of the 5 and 3 mm *accumulated* doses for the spinal cord PRV

BSPRV_acc = Differences between the $D_{0.1cc}$ of the 5 and 3 mm *accumulated* doses for the brainstem PRV

Ips PG_acc = Differences between the D_{mean} of the 5 and 3 mm *accumulated* doses for the ipsilateral parotid gland

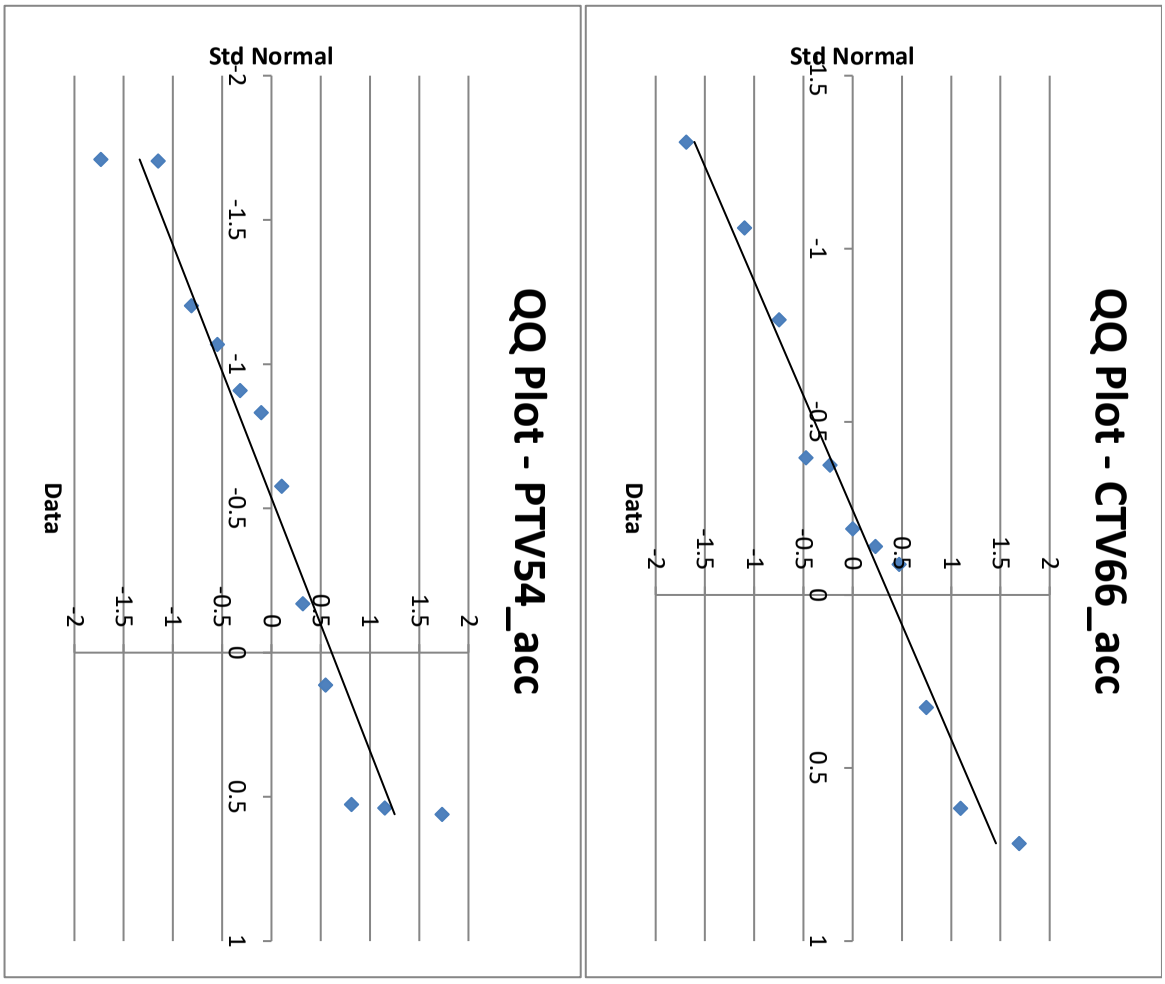
Con PG_acc = Differences between the D_{mean} of the 5 and 3 mm *accumulated* doses for the contralateral parotid gland

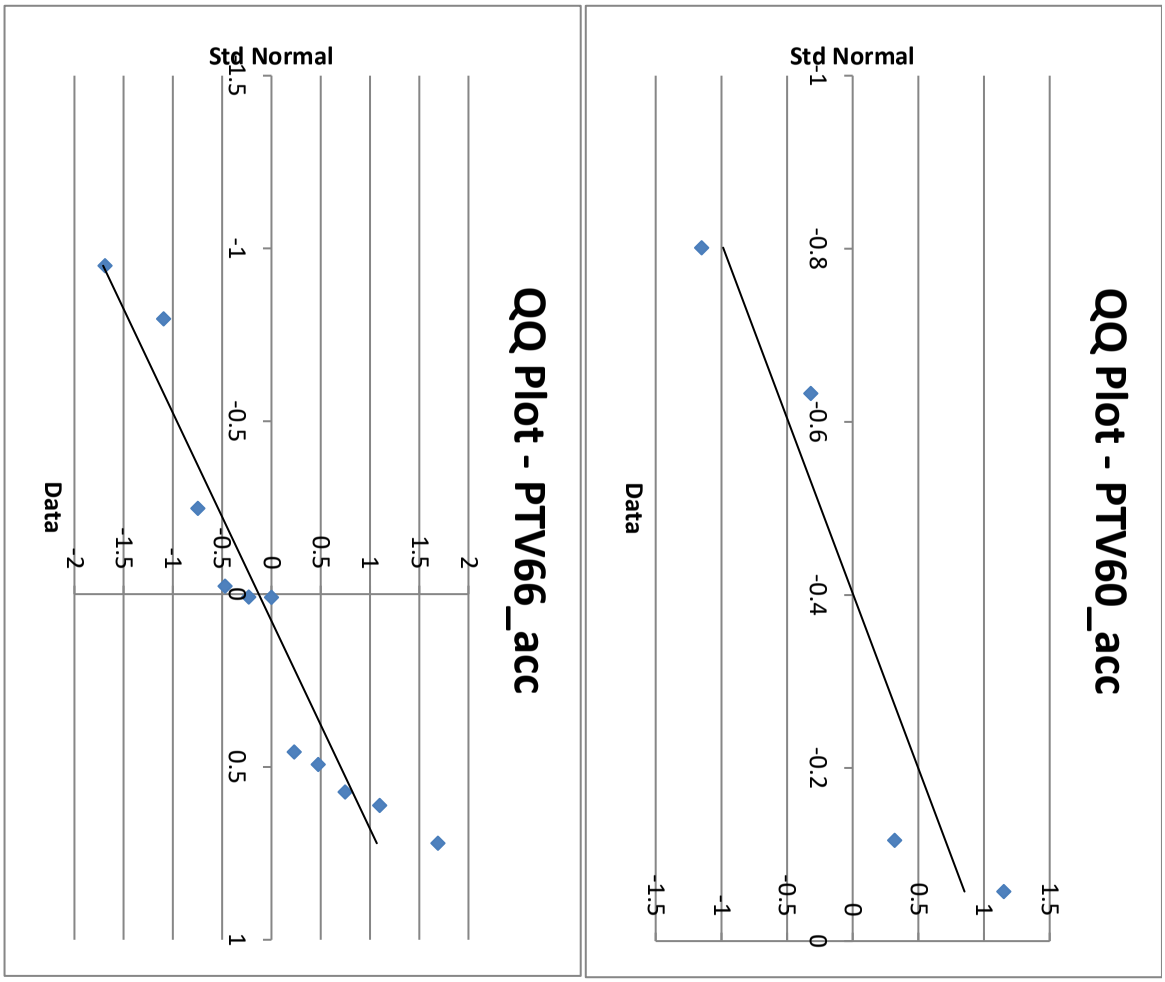
Ips SMG_acc = Differences between the D_{mean} of the 5 and 3 mm *accumulated* doses for the ipsilateral submandibular gland

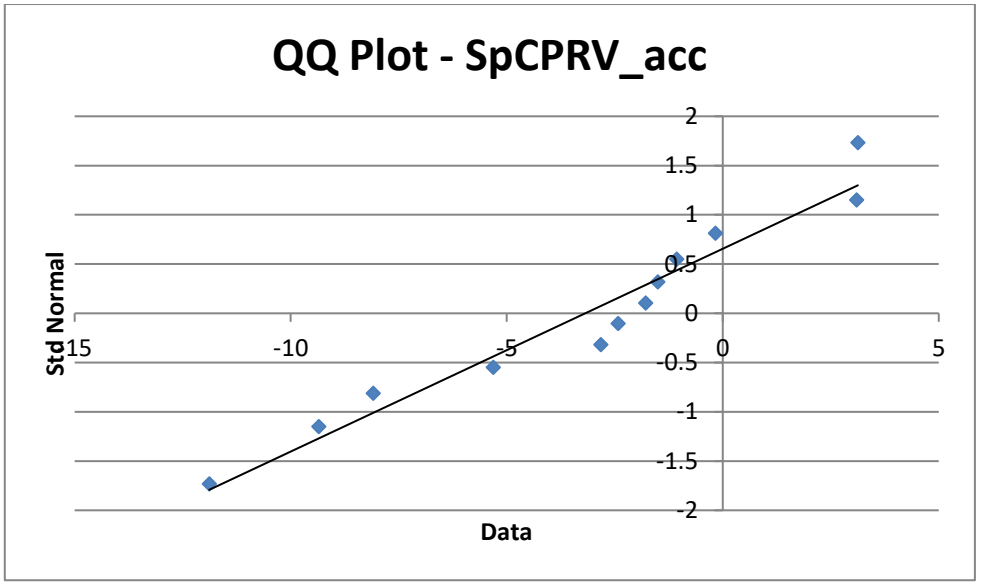
Con SMG_acc = Differences between the D_{mean} of the 5 and 3 mm *accumulated* doses for the contralateral submandibular gland

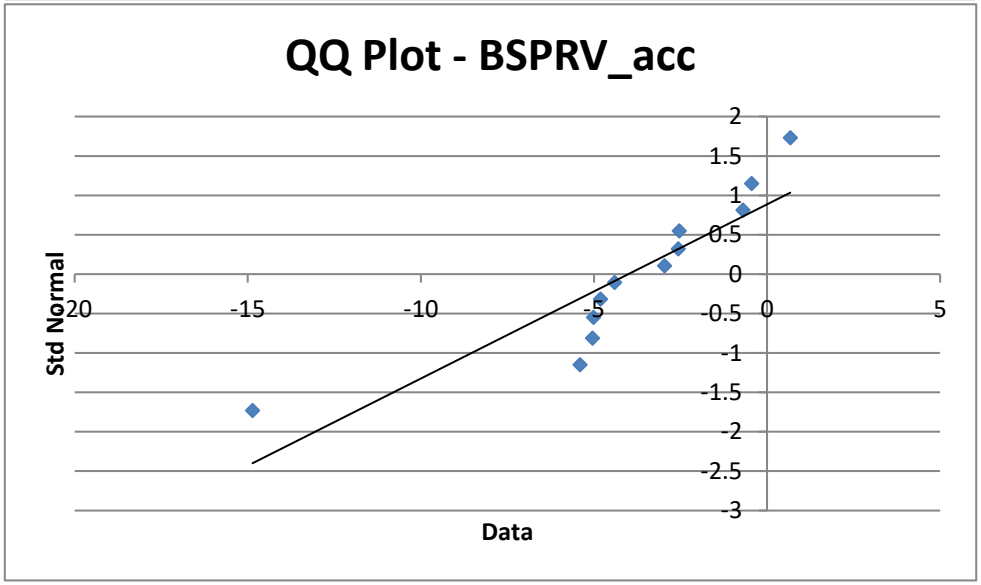
QQ Plot - CTV54_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2
-1.5
-1
-0.5
0
0.5

QQ Plot - CTV60_acc
Std Normal
Data
1.5
1
0.5
0
-0.5
-1
-1.5
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4


QQ Plot - CTV66_acc
Std Normal
Data
QQ Plot - PTV54_acc
Std Normal
Data

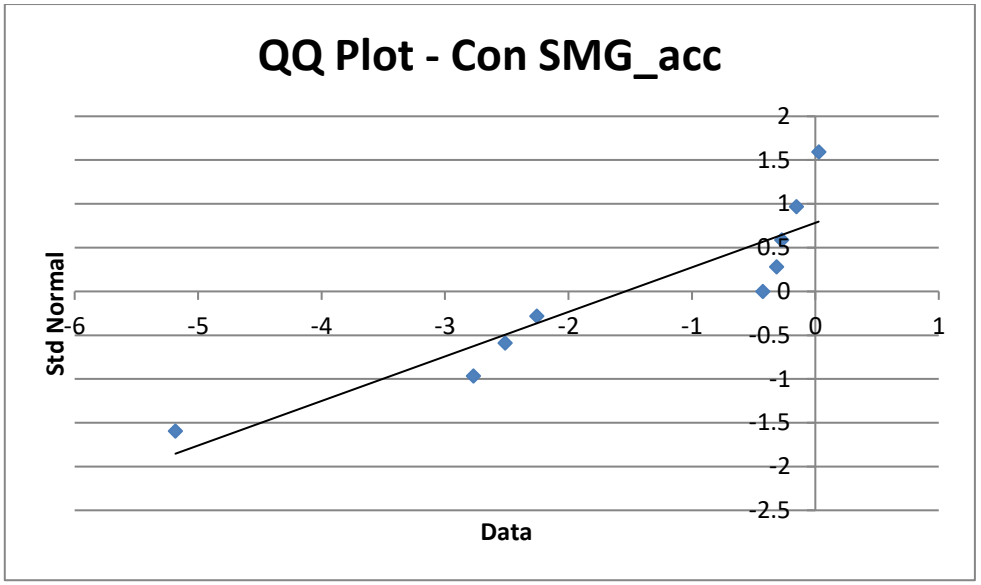

QQ Plot - PTV60_acc
Std Normal
Data
1.5
1
0.5
0
-0.5
-1
-1.5
-1
-0.8
-0.6
-0.4
-0.2
0
QQ Plot - PTV66_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-1.5
-1
-0.5
0
0.5
1


QQ Plot - SpCPRV_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-15
-10
-5
0
5


QQ Plot - BSPRV_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-20
-15
-10
-5
0
5

QQ Plot - Ips PG_acc
Std Normal
Data
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-12
-10
-8
-6
-4
-2
0
2
4

QQ Plot - Con PG_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-6
-5
-4
-3
-2
-1
0


QQ Plot - Con SMG_acc
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-6
-5
-4
-3
-2
-1
0
1

Shapiro-Wilk Test												
	CTV54_dD	*CTV60_dD*	*CTV66_dD*	*PTV54_dD*	*PTV60_dD*	*PTV66_dD*	*pCPRV_dD*	*BSPRV_dD*	*Ips PG_dD*	*Con PG_dD*	*Ips SMG_dD*	*Con SMG_dD*
W-stat	*0.854907*	*0.920262*	*0.937703*	*0.928113*	*0.871635*	*0.859572*	*0.888629*	*0.91328*	*0.912086*	*0.729235*	*0.958837*	*0.924056*
p-value	*0.0*	*0.5*	*0.5*	*0.4*	*0.3*	*0.1*	*0.1*	*0.2*	*0.2*	*0.0*	*0.7*	*0.4*
alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
normal	no	yes	yes	yes	yes	yes	yes	yes	yes	no	yes	yes

CTV_54_dD = Differences between the $D_{99\%}$ of the 5 and 3 mm *ΔD i.e., change in dose* for the LD-CTV

CTV_60_dD = Differences between the $D_{99\%}$ of the 5 and 3 mm *ΔD i.e., change in dose* for the ID-CTV

CTV_66_dD = Differences between the $D_{99\%}$ of the 5 and 3 mm *ΔD i.e., change in dose* for the HD-CTV

PTV_54_dD = Differences between the $D_{98\%}$ of the 5 and 3 mm *ΔD i.e., change in dose* for the LD-PTV

PTV_60_dD = Differences between the $D_{98\%}$ of the 5 and 3 mm *ΔD i.e., change in dose* for the ID-PTV

PTV_66_dD = Differences between the $D_{98\%}$ of the 5 and 3 mm *ΔD i.e., change in dose* for the HD-PTV

SpCPRV_dD = Differences between the $D_{0.1cc}$ of the 5 and 3 mm *ΔD i.e., change in dose* for the spinal cord PRV

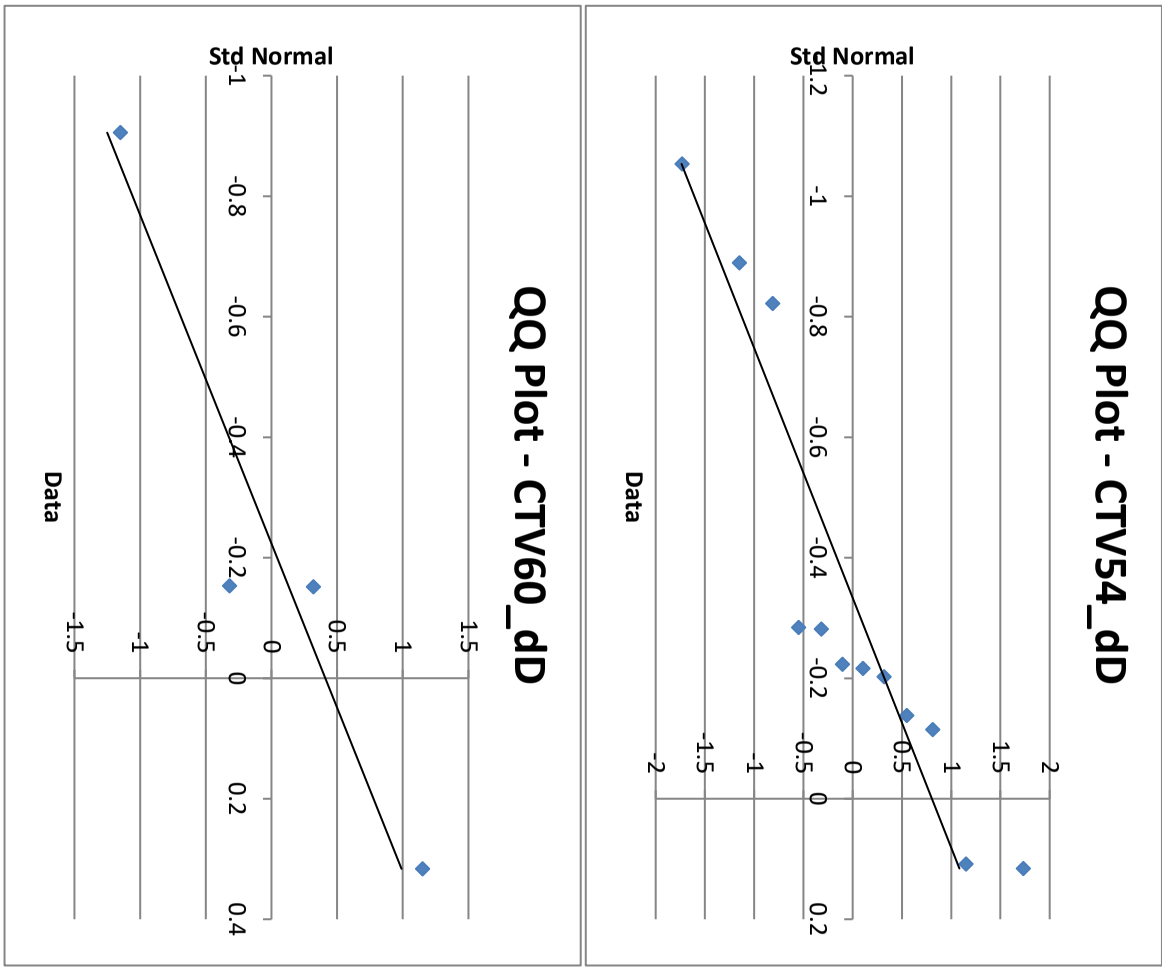
BSPRV_dD = Differences between the $D_{0.1cc}$ of the 5 and 3 mm *ΔD i.e., change in dose* for the brainstem PRV

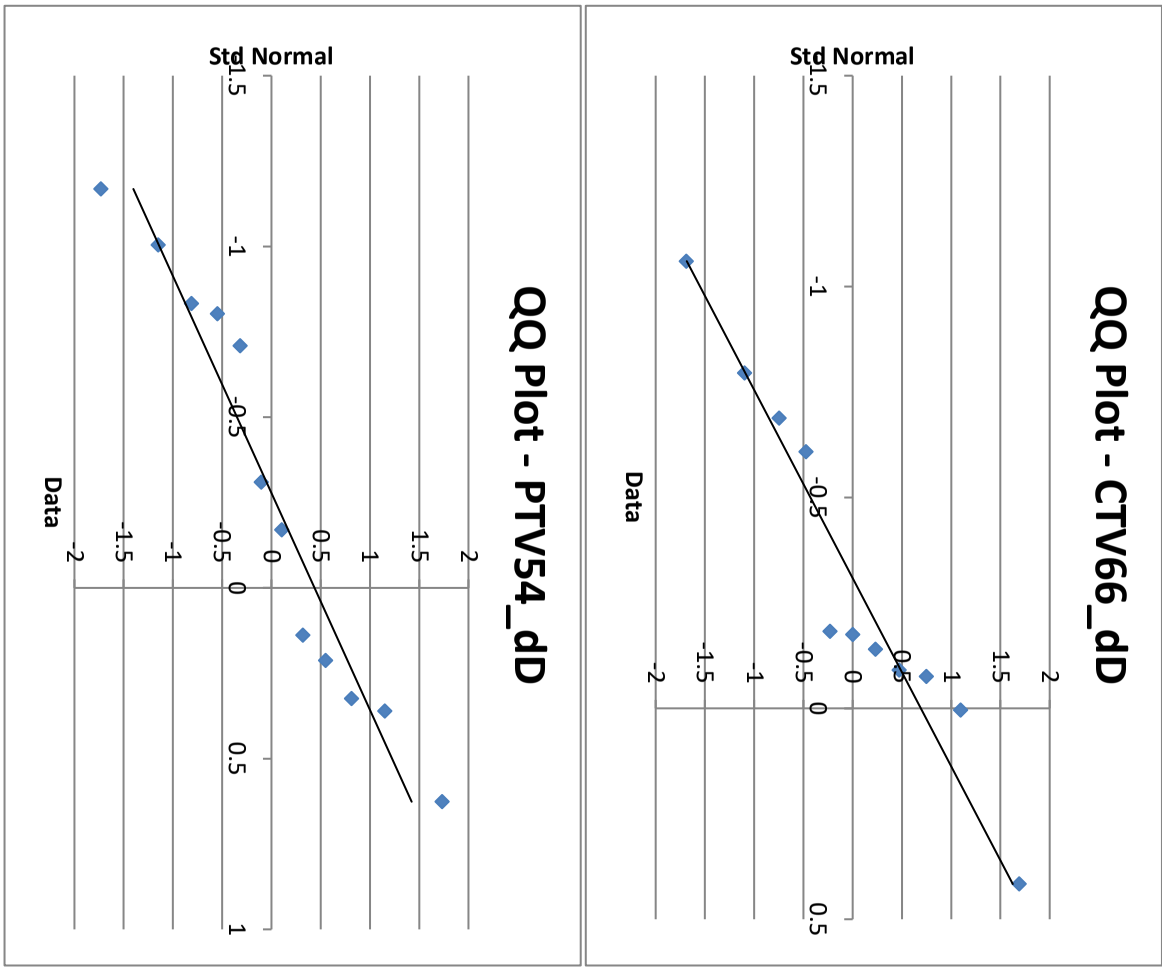
Ips PG_dD = Differences between the D_{mean} of the 5 and 3 mm *ΔD i.e., change in dose* for the ipsilateral parotid gland

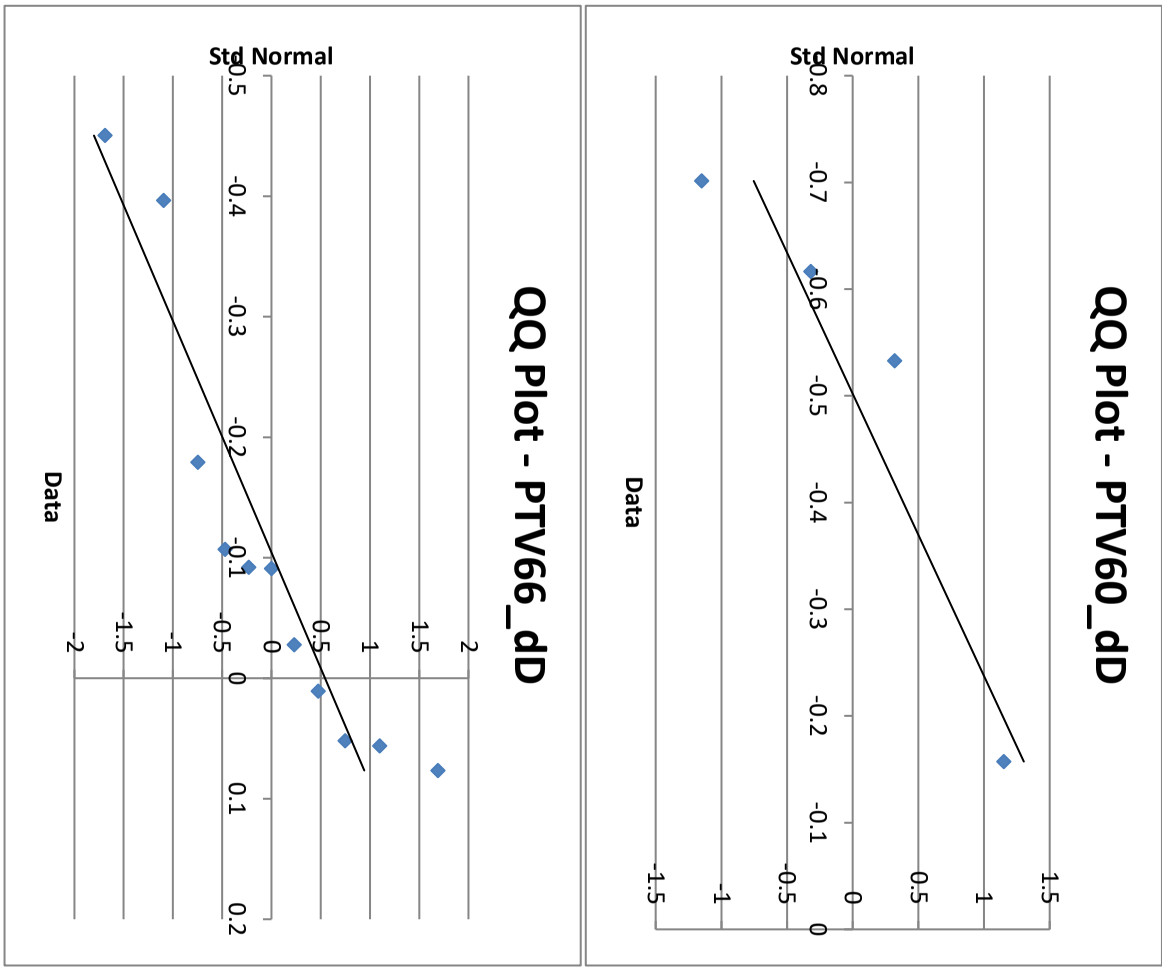
Con PG_dD = Differences between the D_{mean} of the 5 and 3 mm *ΔD i.e., change in dose* for the contralateral parotid gland

Ips SMG_dD = Differences between the D_{mean} of the 5 and 3 mm *ΔD i.e., change in dose* for the ipsilateral submandibular gland

Con SMG_dD = Differences between the D_{mean} of the 5 and 3 mm *ΔD i.e., change in dose* for the contralateral submandibular gland


QQ Plot - CTV54_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-1.2
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
QQ Plot - CTV60_dD
Std Normal
Data
1.5
1
0.5
0
-0.5
-1
-1.5
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4


QQ Plot - CTV66_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
1.5
1
0.5
0
-0.5
-1
-1.5
0.5
1
QQ Plot - PTV54_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
1.5
1
0.5
0
-0.5
-1
-1.5
0.5
1


QQ Plot - PTV60_dD
Std Normal
Data
1.5
1
0.5
0
-0.5
-1
-0.8
-0.7
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0
QQ Plot - PTV66_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2

QQ Plot - SpCPRV_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-2
-1
0
1
2

QQ Plot - BSPRV_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-6
-4
-2
0
2
4
6

QQ Plot - Ips PG_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4

QQ Plot - Con PG_dD
Std Normal
Data
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5

QQ Plot - Ips SMG_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3

QQ Plot - Con SMG_dD
Std Normal
Data
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6

A.17: Individual patient OAR dosimetry

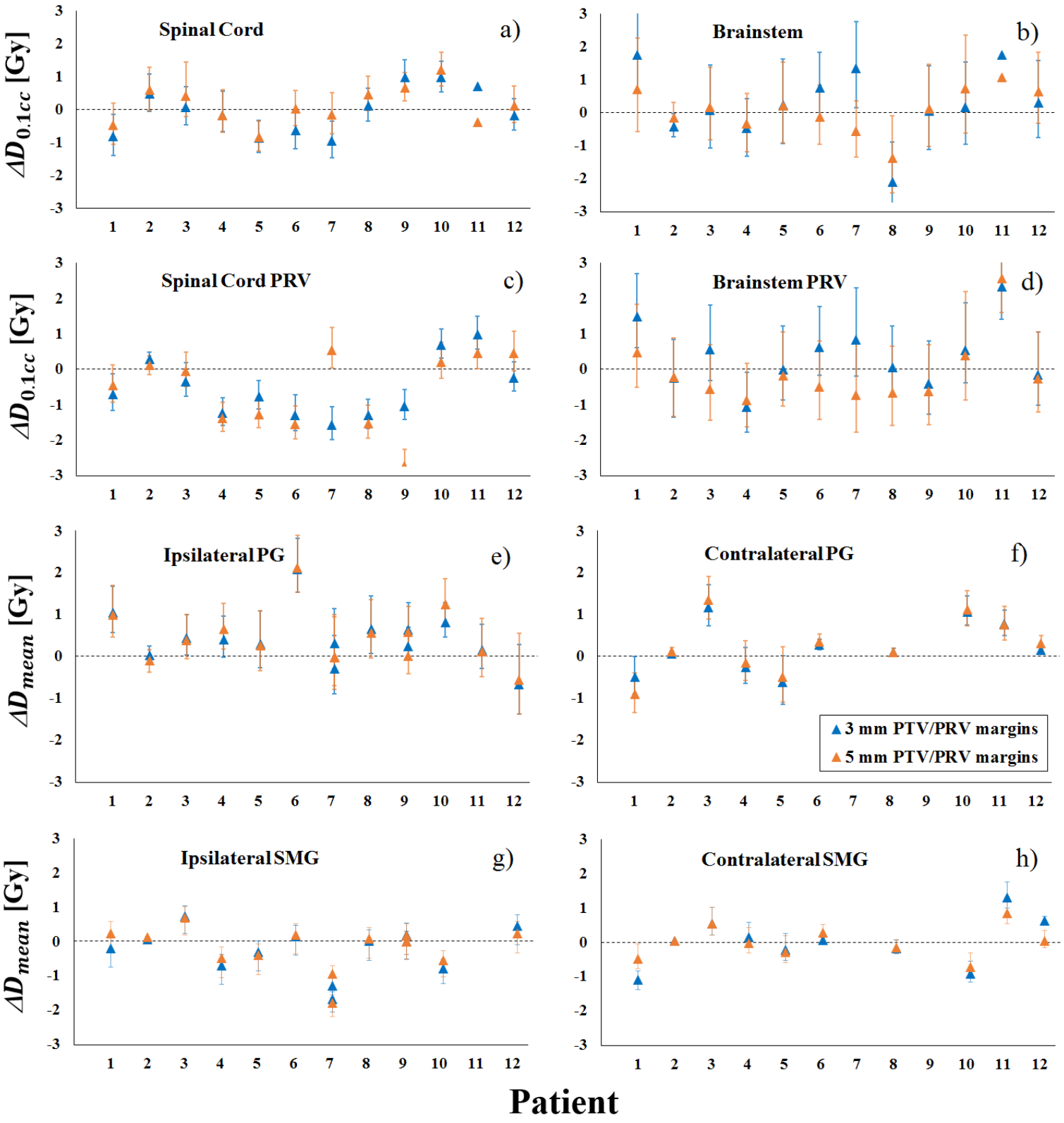
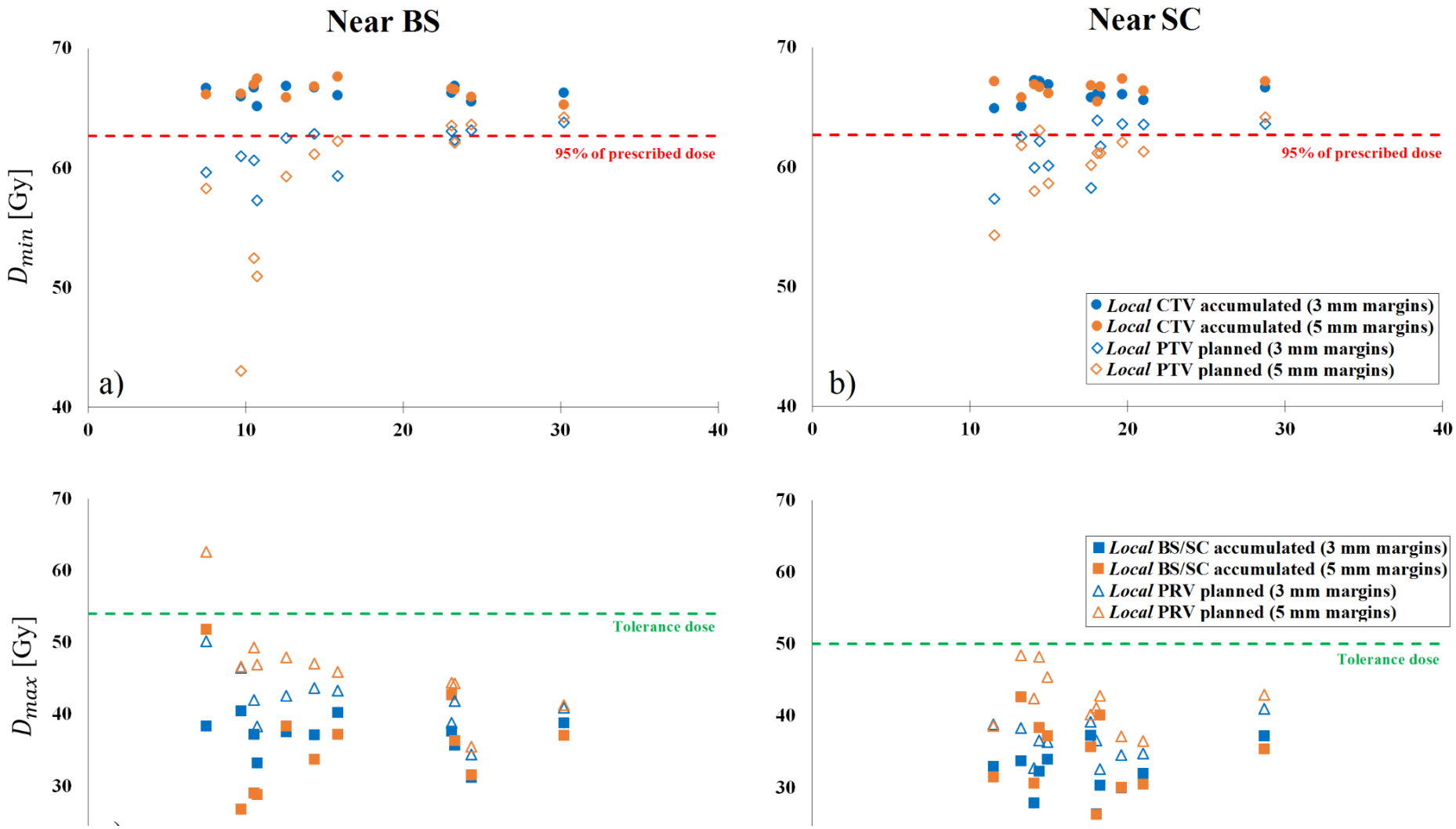


Figure A.17.1: Scatter plots of the planned dose subtracted from the accumulated dose, ΔD for the 12 patients' organs at risk (OARs) when utilizing 3 or 5 mm planning target volume (PTV) and planning risk volume (PRV) margins. PG = parotid gland. SMG = submandibular gland. $\Delta D_{0.1cc}$= minimum dose to 0.1 cm^3 of the volume that receives the highest dose. ΔD_{mean} = mean dose. 95% level of confidence shown by error bars.

Figure A.17.2 shows that a higher D_{min}^{plan} to the *local* PTV (open markers) was achievable with 3 mm compared to 5 mm margin expansions in particular if the minimum distance between HD-CTV and BS was 10 mm or less. There was insufficient data in this range to decide whether this trend was also present for target volumes near the SC (Figure A.17.2b). Figure A.17.2a-b clearly show that the *local* CTV D_{min}^{acc} (solid markers) was well above 95% of D^{presc} for both locations irrespective of the achieved local minimum PTV dose for both margin expansions. Very similar results can be observed for the maximum dose to the OARs and corresponding PRVs (Figure A.17.2c-d). In one case (patient ten), a compromise to the local BS PRV D_{max}^{plan} (62.6 Gy; tolerance dose 54 Gy) was required for the 5 mm expansion to achieve sufficient PTV coverage during treatment planning. It should be reiterated that this compromise was not necessary for the 3 mm margin expansion and that D_{max}^{acc} to the actual BS was within tolerance for both margin expansions (Figure A.17.2c).



A.17.2: The minimum planned D_{min}^{plan} and accumulated dose D_{min}^{acc} for the local PTV and CTV respectively using 3 and 5 mm planning target volume (PTV)/planning risk volume (PRV) margin plans, plotted as a function of the minimum distance (d_{min}) between the clinical target volume (CTV) and brainstem (BS) (a) and spinal cord (SC) (b). The maximum planned D_{max}^{plan} and accumulated dose D_{max}^{acc} for the local PRV and BS/SC respectively using 3 and 5 mm PTV/PRV margin plans, plotted as a function of the minimum distance (d_{min}) between the CTV and BS (c) and SC (d).

A.18: Time-resolved point dose QC analysis

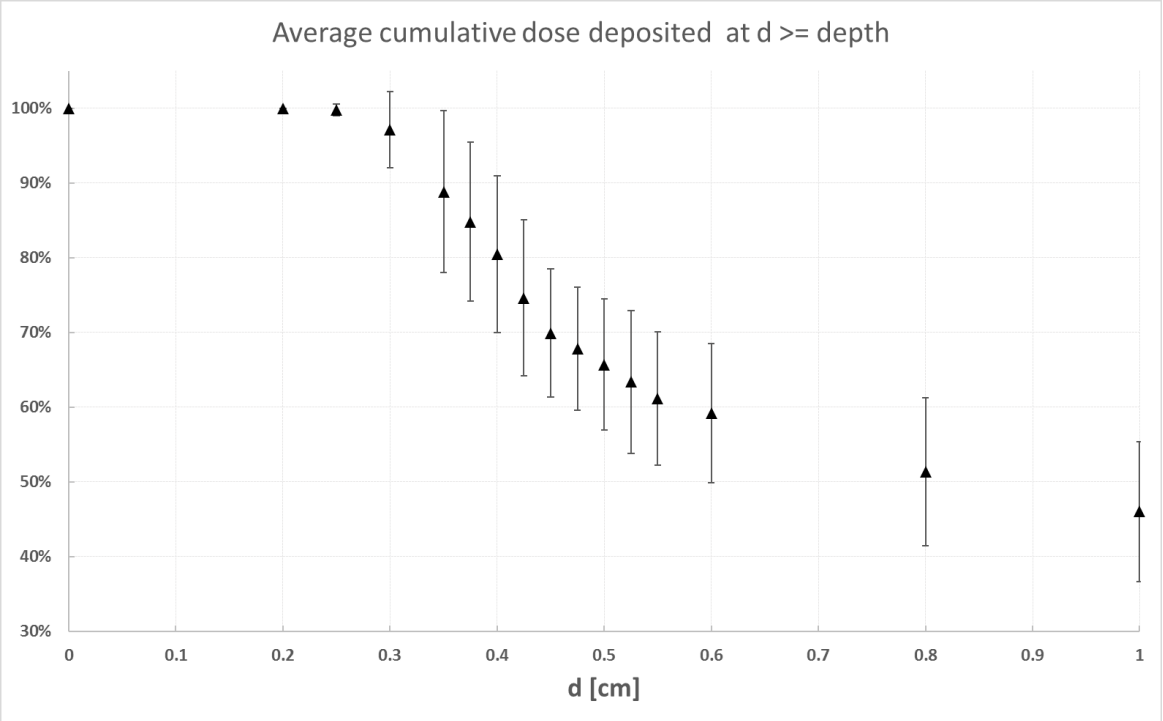


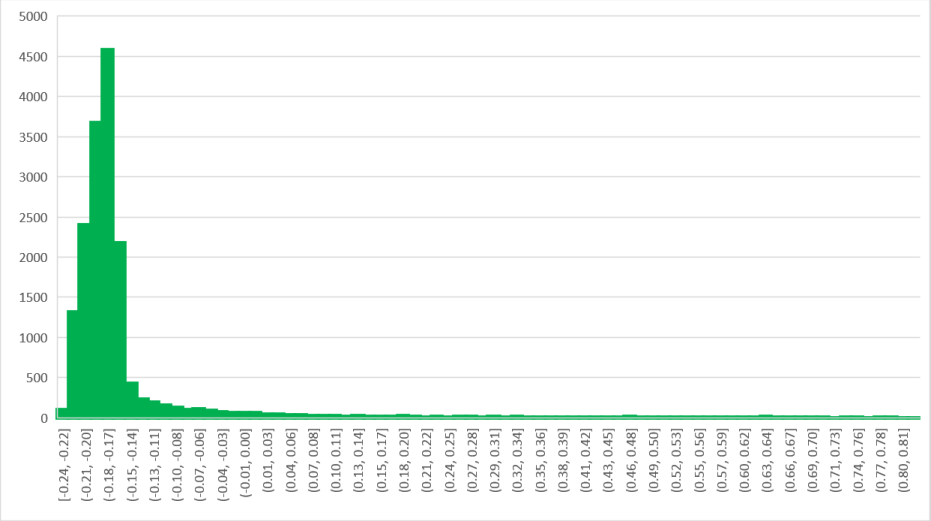
Figure G.1: Average percentage dose deposited at VMAT control points where effective depth (depending on beam angle) was larger than d for 40 points within the superficial aspect of the HD-CTV where the accumulated dose was significantly less than 95% D^{prescr}. The error bars represent +/- 1 S.D.

A.19: HU override analysis

Results of TPS calculations using a HU override equal to +/- 100 HU of the most superficial 4 mm within the body contour. The analysis was conducted for the four patients where a HD-CTV superficial underdosage was observed. First, the voxel-specific dose difference between the +/- 100 HU for each voxel of the HD-CTV was recorded. Second, the 95% percentile range of the voxel-specific dose differences was used to estimate the uncertainty in the dose calculation due to potential uncertainty in the HUs. The largest 95% percentile range (patient 12; 0.68 Gy; ~1% D^{presc}) was used in the statement in the Discussion. The following histograms are of the +/- 100 HU dose difference for the HD-CTVs' voxels.

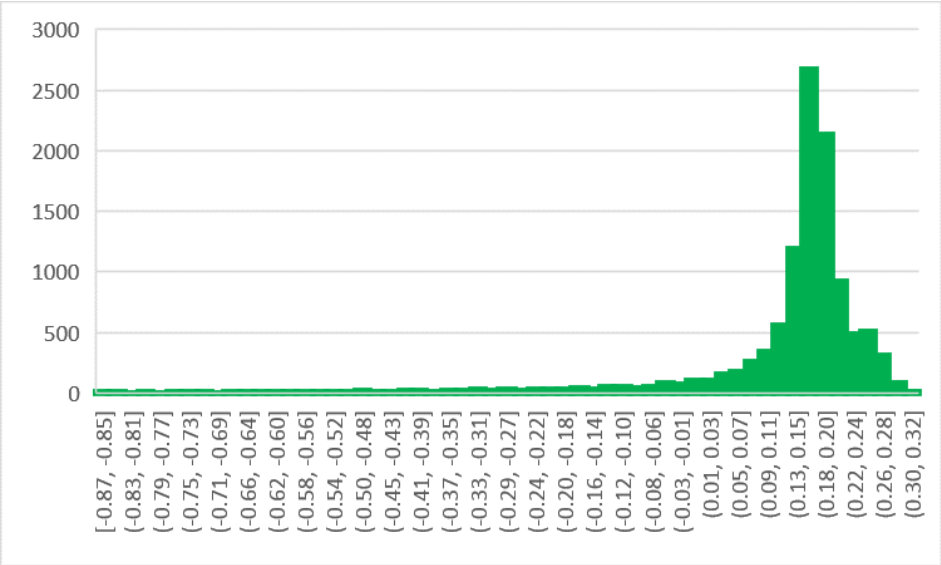
Patient 1

(-100HU to +100HU)	
95% percentile range	
2.50%	97.50%
-0.22	0.15



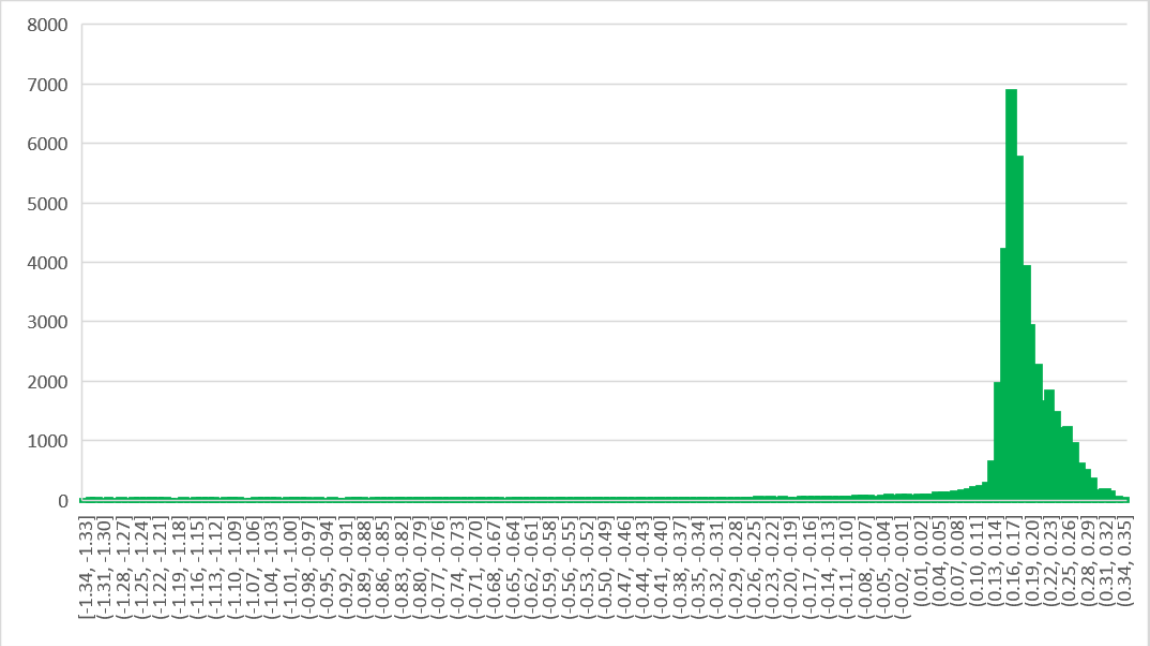
Patient 10

(-100HU to +100HU)	
95% percentile range	
2.50%	97.50%
-0.28	0.27



Patient 11

(-100HU to +100HU)	
95% percentile range	
2.50%	97.50%
-0.05	0.28



Patient 12

(-100HU to +100HU)	
95% percentile range	
2.50%	97.50%
-0.47	0.21

APPENDIX B: CONFERENCE ABSTRACTS

B.1: Deformable image registration of planning CT-to-CBCT for quantifying patient-specific anatomical change during head-and-neck radiotherapy in the context of reduced PTV margins

N Lowther
Wellington Blood and Cancer Centre, Wellington, NZ
nick.lowther@ccdhb.org.nz

D Hamilton
Wellington Blood and Cancer Centre, Wellington, NZ
david.hamilton@ccdhb.org.nz

H Kim
Wellington Blood and Cancer Centre, Wellington, NZ
han.kim@ccdhb.org.nz

J Evans
Wellington Blood and Cancer Centre, Wellington, NZ
jamie.evans@ccdhb.org.nz

S Marsh
University of Canterbury, Department of Physics and Astronomy, Christchurch, New Zealand
steven.marsh@canterbury.ac.nz

R Louwe
Wellington Blood and Cancer Centre, Wellington, NZ
rob.louwe@ccdhb.org.nz

Introduction Reduced treatment toxicity while maintaining equivalent local-regional control rates have been reported when reducing PTVs from 5 to 3 mm for Head and Neck radiation oncology [1,2]. However, many treatment aspects should be carefully considered before the PTV margin can be amended. This study investigates the application of deformable image registration (DIR) and Statistical Process Control (SPC) to quantify deformation, i.e., changes in patient pose and anatomy during treatment as a first step in developing a safety net for PTV margin reduction.

Method This retrospective study includes 12 patients who received planning computed tomography (pCT) and daily CBCTs to verify their positioning. DIR of the pCT-to-CBCTs was performed and the relative position of nine bony anatomy (BA) structures of the head and neck region and the volume of a pre-defined neck region V_n^{DIR} were investigated. The DIR results were validated against manual registration of the BA structures and assessment of a neck volume. Exponentially Weighted Moving Average (EWMA) SPC charts were used to investigate trends in patient positioning accuracy and volume change of the neck. Future results to be included in this study will analyse trends of parotid gland changes using SPC, validated against clinician review.

Results Preliminary results showed that the overall mean difference between DIR and manual registration was 0.0 mm with a random variation and systematic dispersion of 1.0 and 0.4 mm (1 S.D.), respectively. Linear regression showed a high correlation between V_n^{DIR} and a manually contoured volume of a different neck region (r^2=0.849; p<0.001), indicating that V_n^{DIR} is a good indicator of treatment volume changes. EWMA charts were able to detect small shifts in patient positioning accuracy and anatomy during treatment (Figure 1).

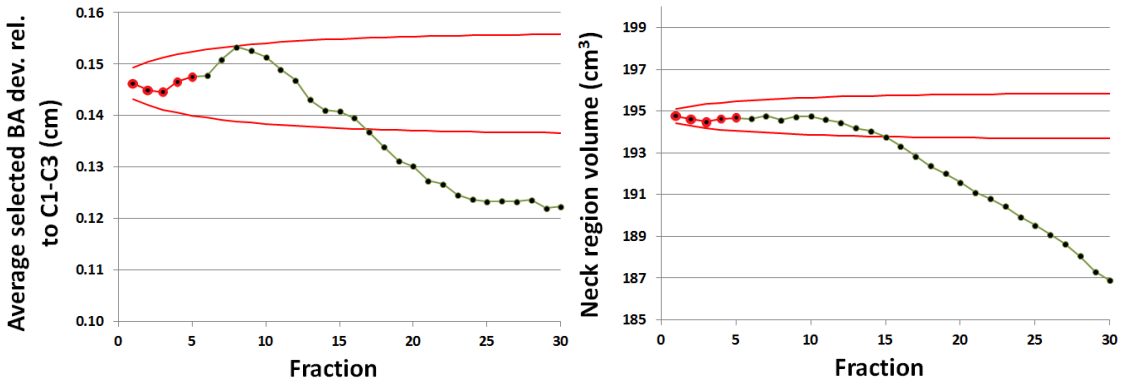


***Figure 1**: EWMA charts for a single patient. The five fraction reference period is shown by red filled circles. Black filled circles indicate the EWMA statistic for each fraction and the red lines indicate upper and lower process limits. (Left) EWMA chart for the average position deviation of selected BA structures relative to C1-C3, showing that patient pose became significantly different from the reference period at fraction 17 (Right) EWMA chart for the volume change of a pre-defined neck region, showing that the volume became significantly different from the reference period at fraction 15.*

Conclusion DIR and SPC provided an effective tool to quantify changes in patient pose and anatomy that may be clinically relevant when reduced PTV margins are used.

B.2: Quantifying anatomical changes using deformable image registration during head-and-neck radiotherapy to facilitate safe PTV margin reduction

N Lowther[1,2], D Hamilton[1], H Kim[1], J Evans[1], S Marsh[2], R Louwe[1]

[1]Wellington Blood and Cancer Centre, Wellington, NZ; [2]University of Canterbury, Department of Physics and Astronomy, Christchurch, NZ

Contact**:** Nick Lowther, WBCC, nick.lowther@ccdhb.org.nz

Introduction: Reduced treatment toxicity while maintaining equivalent local-regional control rates have been reported when reducing PTV margins from 5 to 3 mm for Head and Neck radiation oncology [1,2]. However, many aspects of the departmental treatment accuracy should be carefully considered before adopting this PTV margin reduction. This study investigates the application of deformable image registration (DIR) and Statistical Process Control (SPC) to quantify deformation, i.e., changes in patient pose and anatomy during treatment as a first step in developing a safety net for PTV margin reduction.

Method: This retrospective study included 12 patients who received computed tomography for treatment planning (pCT) and daily CBCTs to verify their treatment position. DIR of the pCT-to-CBCTs was performed to assess changes in the relative position of nine bony anatomy (BA) structures, the volume of a pre-defined neck region V_n^{DIR}, and the relative position and volume of parotid glands during the course of treatment. The DIR results were validated against a manual method of registration (MMR) of the BA structures, independent assessment of a neck volume and clinician re-contour of the parotid glands, respectively. Exponentially Weighted Moving Average (EWMA) SPC charts were used to investigate trends in patient positioning accuracy, volume change of the treatment region and shrinkage and shift of the parotids.

Results: EWMA charts were able to detect trends less than 1 mm in patient positioning reproducibility for individual patients. DIR and MMR results showed an equivalent change in patient pose as a function of time for all BA regions. However for most patients, the systematic difference between the DIR and MMR results for the maxilla and mandible was larger than the uncertainty intervals although they were small compared to the absolute offset of each trend (Figure 1). Weight and parotid gland change trends departed control limits after 5-15 fractions for 50% and 70% of cases, respectively. Future work will correlate these trends to help understand underlying mechanics of the anatomy changes.

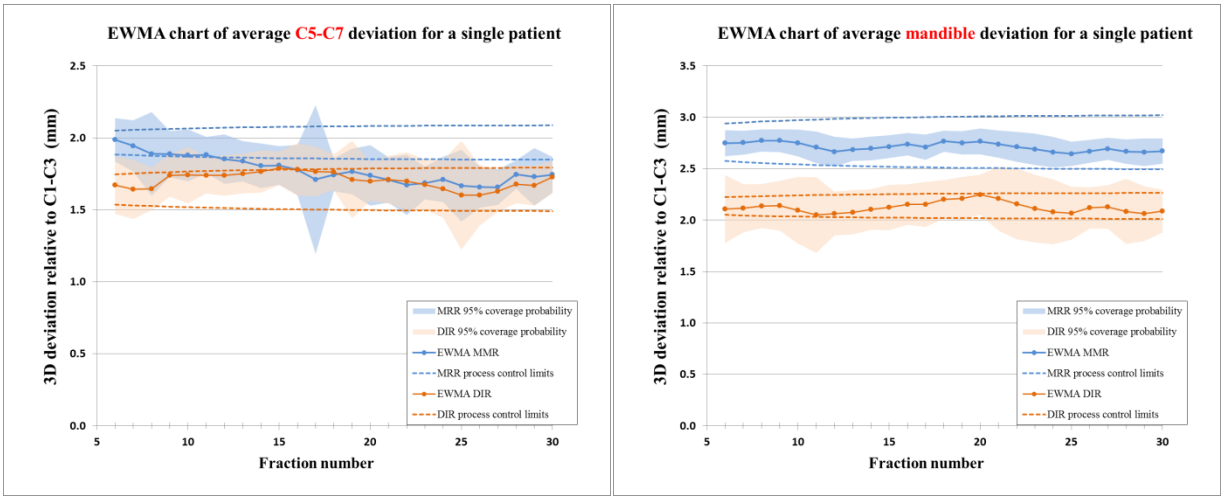


***Figure 1**: EWMA charts of BA deviation for a single patient. Orange and blue filled circles indicate the EWMA statistic for each fraction and the orange and blue dashed lines indicate upper and lower process limits for DIR and MMR, respectively. (Left) EWMA chart for C5-C7 deviation relative to C1-C3. (Right) EWMA chart for mandible deviation relative to C1-C3.*

Conclusion: DIR and SPC were able to monitor changes in patient pose and anatomy during treatment that may be clinically relevant when using reduced PTV margins

B.3: Daily dose accumulation for individual patients to assess the validity of reduced PTV margin plans in head and neck radiation oncology

N Lowther
Wellington Blood and Cancer Centre, Wellington, NZ
nick.lowther@ccdhb.org.nz

K Peick
Wellington Blood and Cancer Centre, Wellington, NZ
karsten.peick@ccdhb.org.nz

S Marsh
University of Canterbury, Department of Physics and Astronomy, Christchurch, New Zealand
steven.marsh@canterbury.ac.nz

R Louwe
Wellington Blood and Cancer Centre, Wellington, NZ
rob.louwe@ccdhb.org.nz

Introduction Reduced treatment toxicity while maintaining equivalent local-regional control rates have been reported after reducing planning target volume (PTV) margins from 5 to 3 mm for Head and Neck radiation oncology [1,2]. This retrospective study investigates the feasibility of a 3 mm PTV margin plan in combination with a 5 mm backup plan to account for clinically relevant anatomical changes. Specifically, the robustness of treatment plans using a 3 or 5 mm PTV margin for anatomical changes will be investigated as well as strategies for timely intervention.

Method Volumetrically modulated arc therapy (VMAT) plans for 12 patients using 3 or 5 mm PTV margins were created. Deformable image registration (DIR) of the planning CT-to-cone beam CTs was performed to reconstruct the daily delivered dose to target and organ-at-risk (OAR) structures. The cumulative DVH was calculated for clinical target volumes (CTVs), PTVs and OARs as input data to develop different strategies for treatment adaptation.

Results Preliminary results for the 3 patients (7 CTVs and 7 PTVs total) demonstrating the largest anatomical changes during treatment showed an average salivary gland sparing of 2.6 Gy (D_{mean}) for 3 mm PTV margin plans. Overall, the increase in reconstructed OAR delivered dose from the start to end of treatment was small. The maximum OAR increase was 2.2 and 2.8 Gy for 3 mm and 5 mm plans for the parotid glands, respectively. For a 3 mm PTV margin, the cumulative D_{98} was less than 95% for 1 CTV (Figure 1) and 6 PTVs. For a 5 mm expansion, this was observed for 6 PTVs but not for the CTVs.

Conclusion PTV margin reduction resulted in increased OAR sparing. Timely detection of inadequate target coverage due to anatomical changes is feasible and enables timely intervention by using a 5 mm PTV margin plan for the remaining treatment fractions.

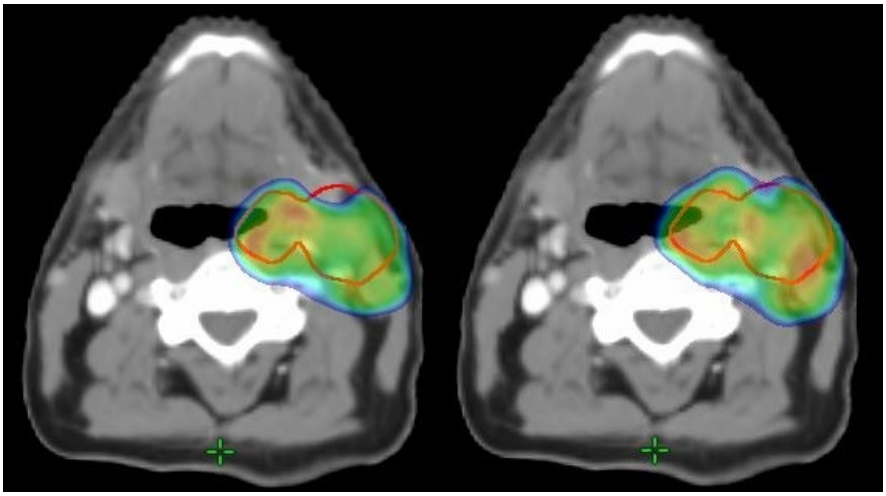


Figure 1: *Deformed high-dose CTV (red contour) and the dose ≥ 95% $D_{prescribed}$ shown as colourwash for patient 1, treatment fraction 20. The CTV coverage (V95%) is less than 98% for the 3 mm PTV margin plan (left) but fully preserved for the 5 mm PTV margin plan (right).*

B.4: Quantifying the dose accumulation uncertainty after deformable image registration to facilitate safe PTV margin reduction in head-and-neck radiotherapy

N Lowther[1,2], S Marsh[2], R Louwe[1]

[1]Wellington Blood and Cancer Centre, Wellington, NZ; [2]University of Canterbury, School of Physical and Chemical Sciences, Christchurch, NZ

Contact: Nick Lowther, WBCC, nick.lowther@ccdhb.org.nz

Introduction: Deformable image registration (DIR) facilitated dose reconstruction is often applied to assess accumulated dose to target volumes and organs at risk (OAR) in the presence of anatomical changes. However, uncertainties in the accumulated dose are seldom quantified. This retrospective study used an *in silico* model based on clinically observed deformations to investigate the uncertainty of reconstructed dose. This verification of the delivered dose may facilitate planning target volume (PTV) margin reduction.

Methods: A cone beam CT (CBCT) that was acquired during the first week of treatment and accurately matched the patient's anatomy of the planning CT (pCT) was selected for 12 patients. These CBCTs were used to generate reference displacement fields, $DF_{B-spline}$ as well as a reference dCBCT^{ref} using B-spline DIR, for 3 treatment fractions. This served as ground truth to assess the accuracy of Varian's SmartAdapt which is based on the demons algorithm. A DF_{demons} was generated by registering the pCT to the dCBCT^{ref}, and both the inverse $DF_{B-spline}$ and DF_{demons} were applied to reconstruct and accumulate the delivered dose in the pCT scan. The total uncertainty of the demon's reconstructed dose (U_{demons}) was assessed by including both the voxel-specific dose difference between the two accumulated dose distributions (∂), and the dose error due to the inverse consistency (IC) of $DF_{B-spline}$which was quantified by subsequent application of the forward and inverse $DF_{B-spline}$.

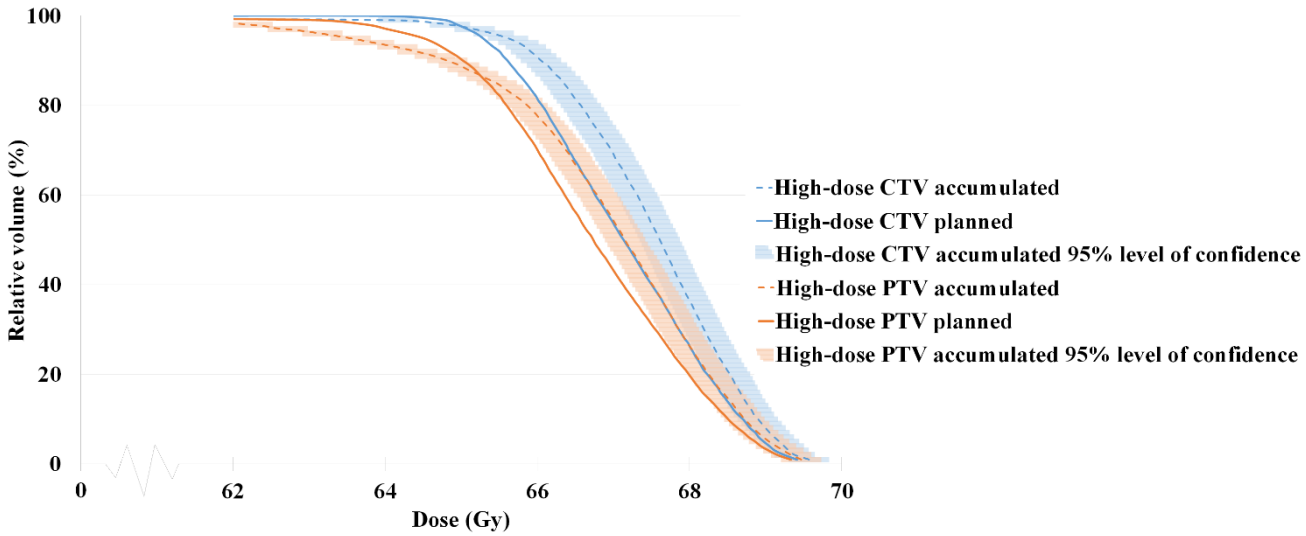


Figure 1: Dose-volume histogram (DVH) of the accumulated dose for a single patient's high-dose clinical target volume (CTV) and planning target volume (PTV) at the end of treatment compared to the original planning DVH.

Results: Analysis of the IC and ∂ distributions showed that U_{demons} can be accurately modeled provided voxels with a net IC displacement larger than a dose calculation voxel were excluded. The total combined uncertainty at the end of treatment was 0.53 Gy at the 95% level of confidence.

Conclusion: The uncertainty in accumulated dose using DIR can be accurately quantified and incorporated in dose-volume histograms. This method can be used to objectively assess the adequacy of target coverage during treatment.

B.5: Statistical process control to monitor anatomical changes during head-and-neck radiotherapy

Nicholas Lowther[1,2], David Hamilton[1], Han Kim[1], Jamie M. Evans[1] Steven Marsh[2] and Rob Louwe[1]
Wellington Blood and Cancer Centre, Wellington, New Zealand[1]
Department of Medical Physics, University of Canterbury, Christchurch, New Zealand[2]

Purpose: Reduced toxicity while maintaining loco-regional control rates have been reported after reducing planning target volume (PTV) margins for head-and-neck radiotherapy (HNRT). In this context, quantifying anatomical changes to monitor the patient during treatment is preferred. This retrospective feasibility study investigated the application of deformable image registration (DIR) and Exponentially Weighted Moving Average (EWMA) Statistical Process Control (SPC) charts for this purpose.

Materials and methods: DIR was performed between the computed tomography for treatment planning (pCT) images of 12 patients and their daily on-treatment cone beam computed tomography (CBCT) images to quantify changes in patient pose and anatomy during treatment. EWMA charts were used to investigate trends in patient positioning reproducibility and soft tissue changes of various structures. The 90% confidence limits for both the EWMA trends and the SPC process limits were obtained using a comprehensive uncertainty analysis. These confidence limits were used to confirm whether a trend breached either an SPC limit or an *a priori* set clinical limit of 2 mm at a previous fraction or not.

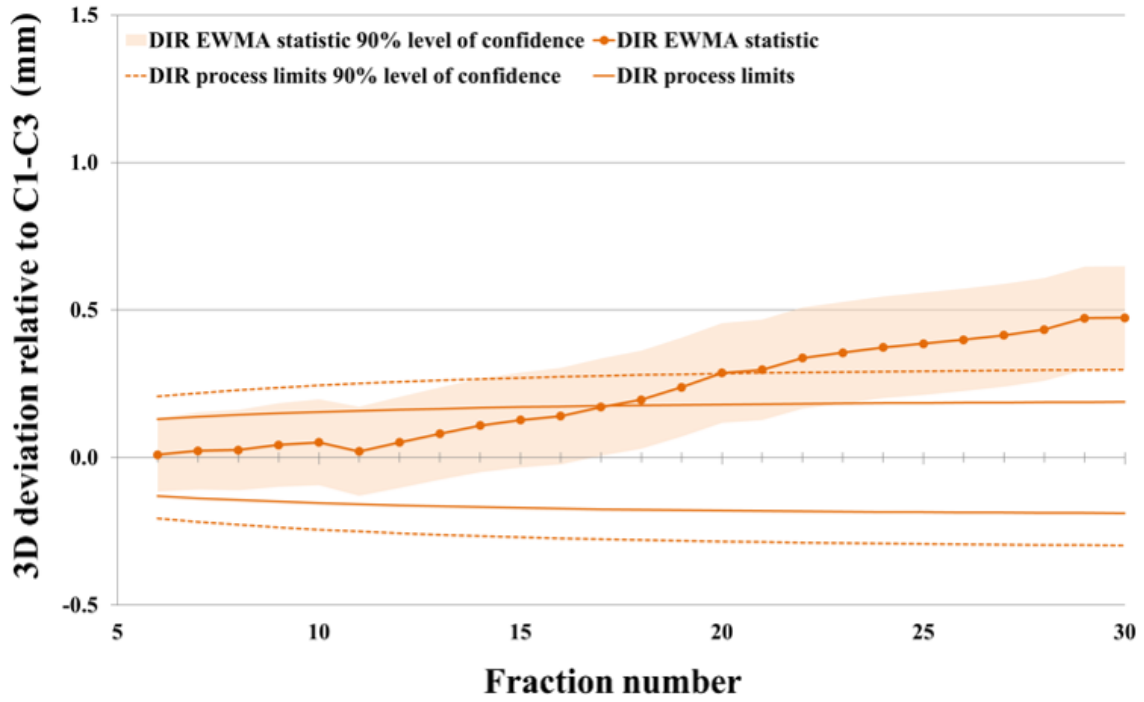


Figure 1: EWMA chart of the mandible deviation for patient 3 relative to week one of treatment

Results:
Trends in patient positioning reproducibility relative to the first week of treatment that were outside SPC process limits before the end of treatment week 4 occurred in 54% of cases. Only 24% of these cases could be confirmed at a 90% confidence level before the end of treatment. Using an *a priori* clinical limit of 2 mm, absolute changes in patient pose were detected in 39% of cases, of which 82%

were confirmed. Soft tissue trends outside SPC process limits occurring before the end of treatment week 4 could be confirmed in 90% of cases.

Conclusion: EWMA trends based on DIR data combined with structure specific action thresholds enabled detection of systematic changes in patient pose and anatomy during the first four weeks of treatment. This approach may facilitate timely treatment adaptation and provide a safety net for PTV margin reduction.

B.6: Dose accumulation assessing the validity of reduced PTV margins in head-and-neck radiotherapy

Nicholas Lowther[1,2], Steven Marsh[2] and Rob Louwe[1]
Wellington Blood and Cancer Centre, Wellington, New Zealand[1]
Department of Medical Physics, University of Canterbury, Christchurch, New Zealand[2]

Purpose: Emerging literature has reported reduced treatment toxicity while maintaining equivalent local-regional control rates after reducing planning target volume (PTV) margins from 5 to 3 mm for head-and-neck radiotherapy (HNRT) [1–3]. As we considered whether it was possible to implement reduced 3 mm PTV margins in our department, it was recognized that many aspects of HNRT including robustness of the planning solution for anatomical changes may influence treatment outcomes and should be considered when PTV margins are reduced. This retrospective study investigates the robustness of treatment plans using 3 or 5 mm PTV margins for anatomical changes. The results of this study can be used to develop strategies for treatment adaption based on objective criteria.

Methods: Volumetrically modulated arc therapy (VMAT) plans for 12 patients using 3 or 5 mm PTV margins (Prescribed dose 54 Gy and simultaneous integrated boost volumes to 60 and 66 Gy in 30 fractions) were optimized using the local planning protocol. The planning CT (pCT) was first registered to each daily cone beam CT using deformable image registration (DIR). Subsequently, the inverse registration was used to reconstruct and accumulate the delivered dose to target and organ-at-risk (OAR) structures in the pCT scan. For the initial analysis, the coverage of the PTVs, clinical target volumes (CTVs) and salivary glands were assessed using the $D_{98\%}$, $D_{99\%}$ and D_{mean} respectively. The uncertainty of the reconstructed dose was assessed using an *in silico* model based on clinically observed deformations to determine the 95% level of confidence.

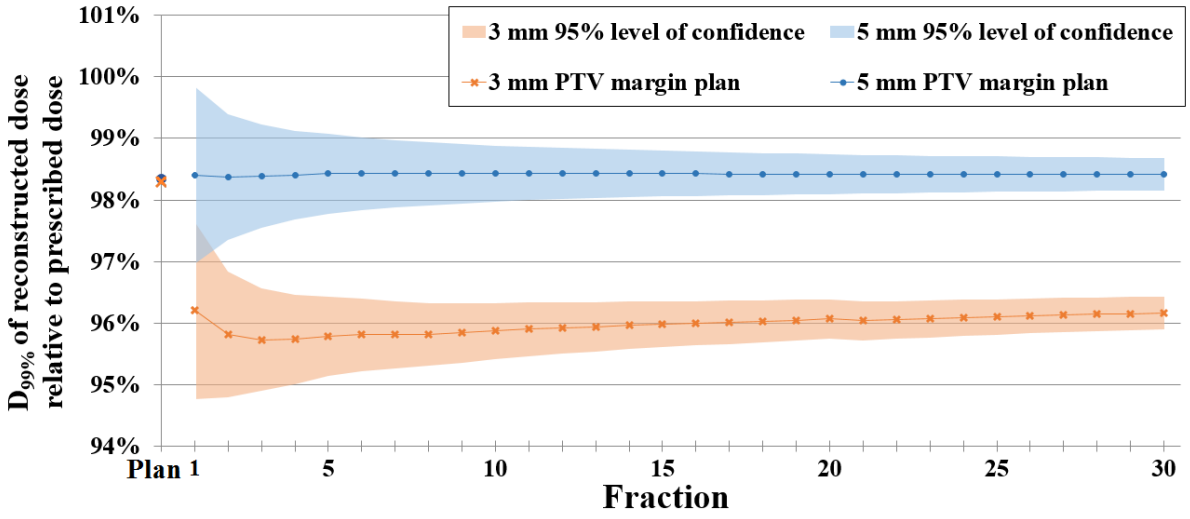


Fig. 1: Reconstructed dose relative to prescribed dose $D_{99\%}$ for an individual patient's high-dose CTV.

Results: Preliminary results (4 patients) showed that for 4 out of 5 high-dose PTVs, $D_{98\%}$ during treatment was at least 2% lower than planned in both margin plans. The high-dose CTVs' $D_{99\%}$ generally changed less than 1% relative to the planned dose. However, for two 3 mm plans the high-dose CTV $D_{99\%}$ was systematically 2% lower (Fig.1). For all high-dose CTVs, the $D_{99\%}$ of the reconstructed dose at the end of treatment was always 96% of the prescribed dose or higher in both 3 and 5 mm margin plans. For the elective low-dose CTVs, the reconstructed $D_{99\%}$ was at least 99% of the prescribed dose for all fractions in both plans. On average, the D_{mean} for the ipsilateral parotid glands was 4 Gy lower (range; 1 to 10 Gy) in 3 mm margin plans compared to the 5 mm plans. For submandibular glands, the average dose reduction was 2 Gy. The difference between reconstructed and planned D_{mean} of the parotid glands was not significantly different for the 3 mm (range; -1.2 to 1.1 Gy) and the 5 mm (range; -1.4 to 1.2 Gy) plans (p=0.67).

Conclusion: Changes in delivered dose due to anatomical changes can be accurately reconstructed using DIR. Initial results indicate that 3 mm PTV margins are robust for anatomical changes occurring during HNRT with CTV $D_{99\%} \geq 96\%$ of the prescribed dose. PTV margin reduction resulted in increased OAR sparing.

B.7: Dose accumulation assessing the validity of reduced PTV margins in head-and-neck radiotherapy

N Lowther
Wellington Blood and Cancer Centre, Wellington, NZ
nick.lowther@ccdhb.org.nz

S Marsh
University of Canterbury, Department of Physics and Astronomy, Christchurch, New Zealand
steven.marsh@canterbury.ac.nz

R Louwe
Wellington Blood and Cancer Centre, Wellington, NZ
rob.louwe@ccdhb.org.nz

Introduction Emerging literature has reported reduced treatment toxicity while maintaining equivalent local-regional control rates after reducing planning target volume (PTV) margins from 5 to 3 mm in head-and-neck radiotherapy (HNRT) [1–3]. This retrospective study utilised dose accumulation to investigate the robustness of 5 and 3 mm PTV margin treatment plans.

Method Volumetrically modulated arc therapy (VMAT) plans for 12 patients using 5 or 3 mm PTV margins were optimised using the local planning protocol. The planning CT (pCT) was first registered to each daily cone beam CT using deformable image registration (DIR). Subsequently, the inverse registration was used to reconstruct and accumulate the delivered dose to the patient in the pCT scan. Accumulated dose-volume histograms (DVHs) were constructed and coverage of the PTVs, clinical target volumes (CTVs) and parotid glands were assessed using the $D_{98\%}$, $D_{99\%}$ and D_{mean}, respectively.

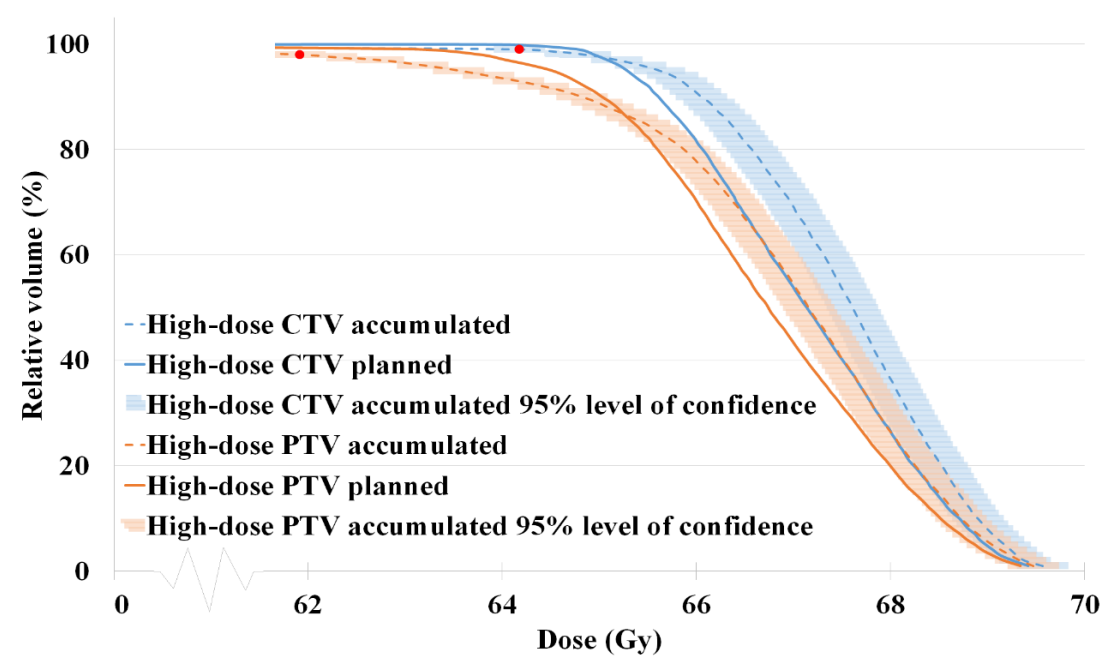


Fig 1*: Dose-volume histogram (DVH) of the accumulated dose for patient 1's high-dose CTV and PTV at the end of treatment compared to the original planning DVH. The accumulated $D_{98\%}$ and $D_{99\%}$ (red bullet points) of the high-dose PTV and CTV were significantly different at the 95% level of confidence from the planned values but still 97% and 99% of the planned D_{98} and D_{99}, respectively.*

Results Anatomical changes occurring during treatment generally impacted the high-dose CTVs' $D_{99\%}$ by less than 1% relative to the *planned* dose. However, for two 3 mm plans and one 5 mm plan the high-dose CTVs' $D_{99\%}$ were systematically 2% lower due to systematic local position errors or a reduction in the external contour. For 10 out of 11 high-dose CTVs,

the $D_{99\%}$ of the accumulated dose at the end of treatment was at least 95% of the *prescribed* dose for both margin plans. The difference between accumulated and planned D_{mean} of the parotid glands was not significantly different for the 3 mm (range; -0.7 to 2.1 Gy) and the 5 mm (range; -0.9 to 2.1 Gy) plans (p=0.58).

Conclusion Patients at risk of CTV coverage $D_{99\%}$ < 95% of the *prescribed* dose can be identified during treatment using dose accumulation. Loss of coverage for high-dose CTVs using 3 mm and 5 mm PTV margin plans occurred with the same frequency.

B.8: DIR based dose accumulation to validate reduced PTV and PRV margins in head-and-neck radiotherapy

Nicholas Lowther[1,2], Steven Marsh[2] and Rob Louwe[1]
Wellington Blood and Cancer Centre, Wellington, New Zealand[1]
School of Physical and Chemical Sciences, University of Canterbury, New Zealand[2]

Purpose or Objective: Emerging literature has reported reduced treatment toxicity in head-and-neck radiotherapy (HNRT) with 3 instead of 5 mm planning target volume (PTV) margins [1–4]. However, the loco-regional (LR) control rate was not preserved in one study [4]. The current study investigates the occurrence, location and trends where clinical target volume (CTV) coverage and critical organ at risk (OAR) sparing may be at risk in the context of reduced PTV and planning risk volume (PRV) margins. It specifically focusses on locations where the high-dose target coverage is commonly tight during treatment planning.

Material and Methods: VMAT plans for 12 patients were optimised using either 3 or 5 mm PTV and PRV margins (prescribed dose, D^{presc} 66 Gy in 30 fractions). The planning CT (pCT) was first registered to each daily CBCT using demons deformable image registration (DIR). Subsequently, the inverse registration was used to reconstruct and accumulate the delivered dose (D^{acc}) over all fractions. The 95% level of confidence of D^{acc} was determined using the structure- and voxel-specific uncertainty as described previously. The CTV and critical OAR doses were assessed using the dose-volume histogram (DVH) metrics $D^{acc}_{99\%}$ and $D^{acc}_{0.1cc}$, respectively. In addition, the planned and accumulated minimum dose D^{plan}_{min} and D^{acc}_{min} were determined for *local* target volumes defined at the point of minimum distance to the brainstem (BS), spinal cord (SC) and skin.

Results: For each critical OAR and for both margin expansions, $D^{acc}_{0.1cc}$ was below the specific OAR tolerance dose. For the CTVs, $D^{acc}_{99\%}$ was less than 95% D^{presc} at the 95% level of confidence in two and one cases for 3 and 5 mm PTV/PRV margin plans, respectively, with the under dosage occurring near the skin. Fig. 1 shows that target volumes closer to the skin generally exhibited a lower D^{plan}_{min} and a larger drop in *local* coverage during treatment compared to less superficial CTVs, independent of the applied PTV/PRV margin. Analysis of the *local* target coverages near the BS showed that for one patient, an increased BS PRV dose could be avoided using 3 mm margins. For cases with more proximal CTVs, a higher D^{plan}_{min} to the *local* PTV was achievable with 3 mm plans. Nevertheless, the drop in coverage during treatment for the *local* CTV was always less than 2%, with D^{acc}_{min} > 95% D^{presc} (Fig 2). Similar results were obtained for CTV coverage near the SC.

Conclusion: For CTVs proximal to BS and SC, 3 mm margin plans with improved PTV coverage and PRV sparing preserved CTV coverage during treatment. Consistent with results of Chen et al. [2] who reported that ~50% of marginal LR recurrences were observed in superficial areas, our study *only* showed a drop in CTV coverage below 95% D^{presc} in skin regions. Careful monitoring of the delivered dose is therefore warranted both for 3 and 5 mm PTV/PRV margins.

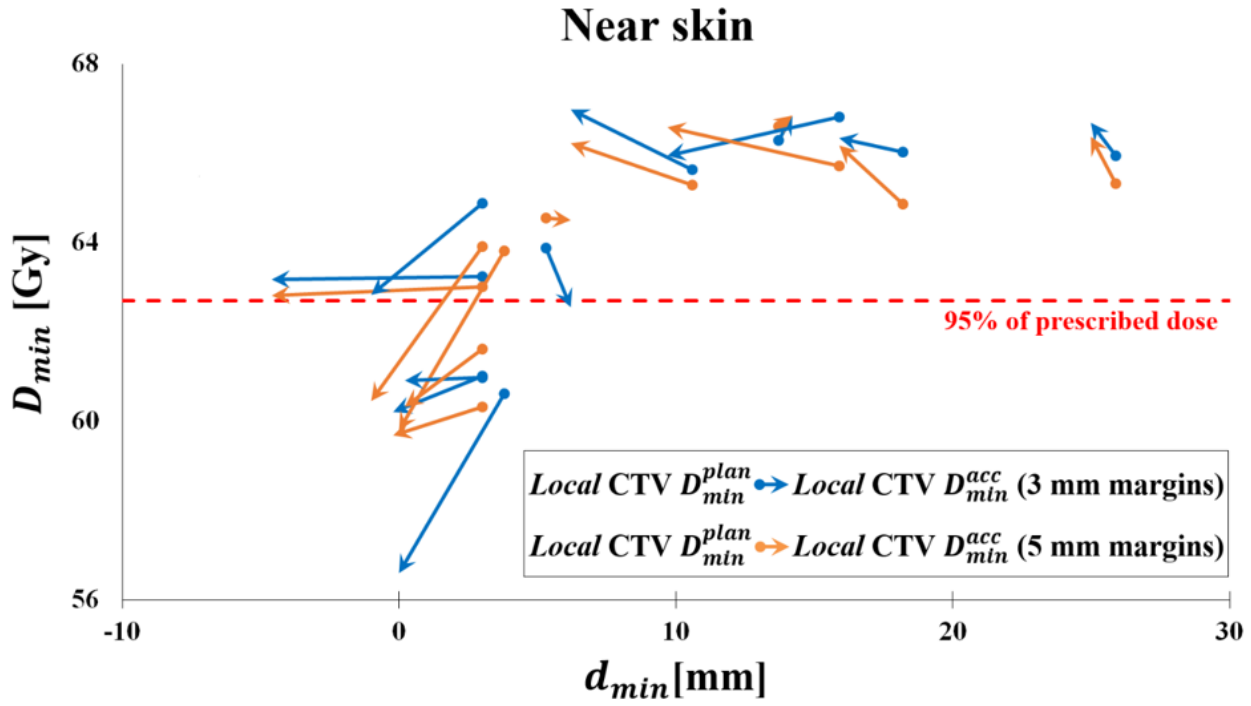


Figure 1: Progression of the minimum planned dose D_{min}^{plan} to the minimum accumulated dose D_{min}^{acc} for local CTVs near the skin for 3 and 5 mm PTV/PRV margin plans. D_{min}^{plan} is plotted as a function of the minimum distance between the CTV and skin. D_{min}^{acc} is plotted as a function of the sum of this minimum distance, and the local shift of the skin at the end of treatment to indicate the observed anatomical changes.

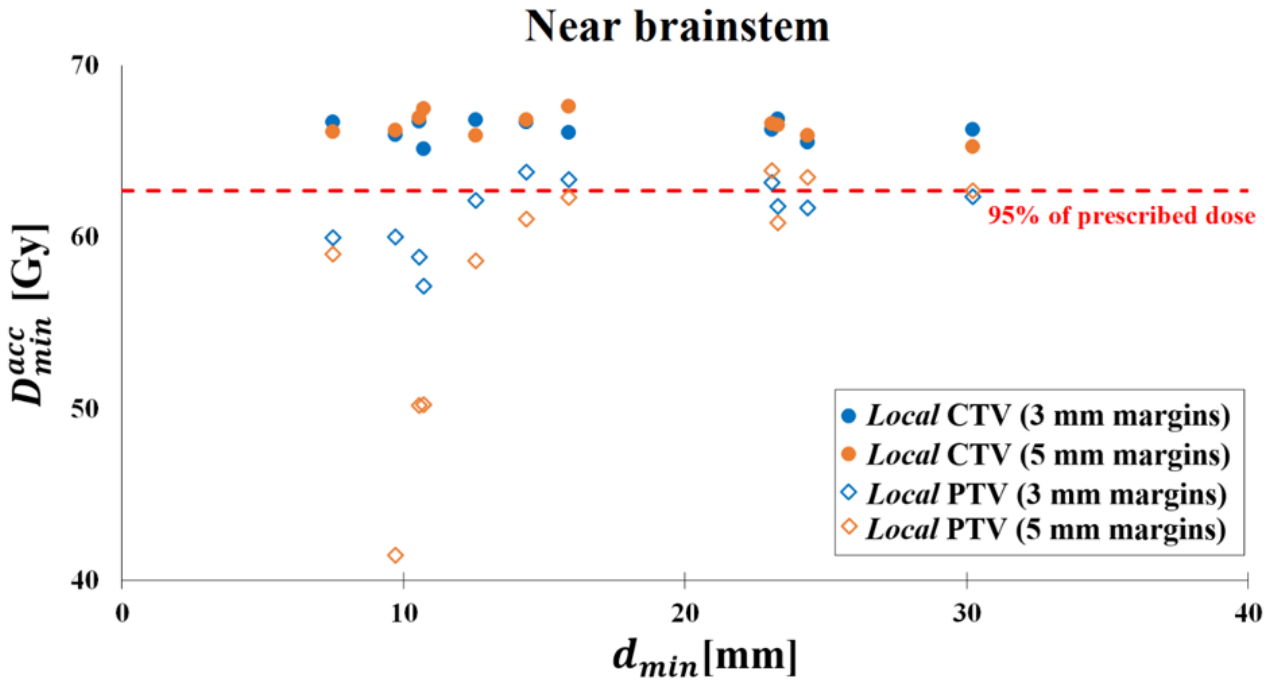


Figure 2: The minimum accumulated dose D_{min}^{acc} for the local PTV and CTV using 3 and 5 mm PTV/PRV margin plans, plotted as a function of the minimum distance (d_{min}) between the CTV and brainstem.

B.9: Quantifying the dose accumulation uncertainty after DIR in head-and-neck radiotherapy

Nicholas Lowther[1,2], Steven Marsh[2] and Rob Louwe[1]
Wellington Blood and Cancer Centre, Wellington, New Zealand[1]
School of Physical and Chemical Sciences, University of Canterbury, New Zealand[2]

Purpose or Objective: Deformable image registration (DIR) facilitated dose reconstruction and accumulation can be applied to assess delivered dose in the presence of anatomical changes and verify the validity of a treatment plan during treatment. However, displacement field errors (DFEs) in the forward registration or inverse consistency (IC) errors in the backward registration introduce uncertainties in the reconstructed dose. This study uses an *in silico* model based on clinically observed deformations as ground truth to investigate the dose accumulation uncertainty (DAU) in head-and-neck radiotherapy (HNRT).

Material and Methods: A planning CT (*pCT*), cone beam CT (CBCT) from week one of treatment and three later CBCTs were used to generate three corresponding *in silico* reference CBCTs serving as ground truth (GT) for 12 HNRT patients. To assess the DAU, the *pCT* was first matched to each GT CBCT using the same reference DIR algorithm (B-spline) to generate a deformed pCT (*dCT*). The treatment plan was then recalculated on the *dCT* and the resulting dose distribution mapped back to the *pCT* space using the backward deformation vector field (DVF) to produce a fraction-specific GT reconstructed dose (RD). The process was repeated using the DIR algorithm under investigation (demons) to generate a demons RD which was compared to the GT RD to calculate the voxel-specific dose error Δ^{tot}. Inverse consistent *c* and inconsistent *i* voxels were identified by successive application of the forward and backward demons DVF, and comparing the net shift to the dose calculation grid size. The fraction-specific RD uncertainties for each structure *S* and both voxel distinctions (i.e., u_c^S and u_i^S) were estimated using the 95% percentile range of Δ^{tot}. The feasibility of incorporating the DAU as a confidence interval in dose-volume histograms (DVHs) of the delivered dose was demonstrated for various structures. For each structure, the voxel-specific uncertainties in the accumulated dose (AD) over all fractions uA^S were first calculated by summing u_c^S or u_i^S in quadrature. Subsequently, the 95% level of confidence for each dose level bin in a DVH of the AD was calculated by averaging the uA^S of all voxels with an AD equal to or larger than that dose level.

Results: Overall, the IC rate of voxels was 98.5% (Fig 1) with u_c^S per fraction equal to [-2.3%; +2.1%], [-10.2%; +15.2%] and [-9.5%; +12.5%] relative to their planned dose for target structures, critical OARs and non-critical OARs, respectively. Inverse *in*consistent voxels generally showed a higher level of uncertainty. At the end of treatment, uA^S was [-0.4%; +0.4%] for target structures. The impact of DAU is demonstrated in the DVH of Fig. 2.

Conclusion: This method can be used in a prospective way to estimate the accuracy of DIR based dose reconstruction and provides a method to objectively assess the adequacy of treatment in the context of anatomical changes or variations in patient positioning. The DAU based on DIR can be accurately quantified and incorporated in DVHs.

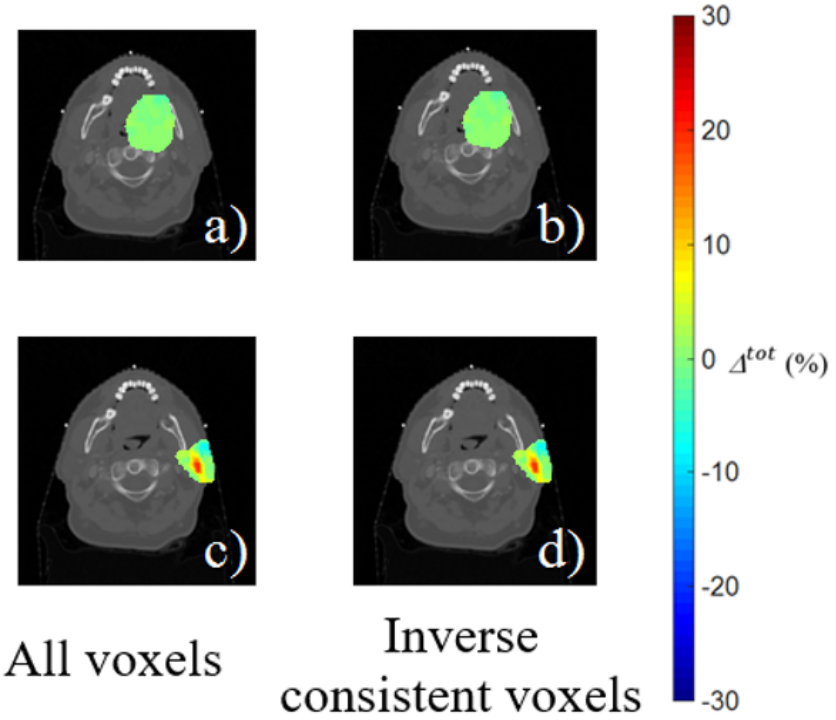


Figure 1: Δ^{tot} of all voxels (left column) and inverse consistent voxels only (right column) for a single fraction. The high-dose PTV (a,b) and ipsilateral parotid gland (c,d) are shown for a single axial slice. Typically, a high IC rate is observed for most structures as can be observed by comparing the left-hand and right-hand columns. The uncertainty in reconstructed dose is generally less than 10% (a & b) but in a small number of cases an uncertainty larger than 10% was observed, as indicated by the region with the red colour wash (c & d).

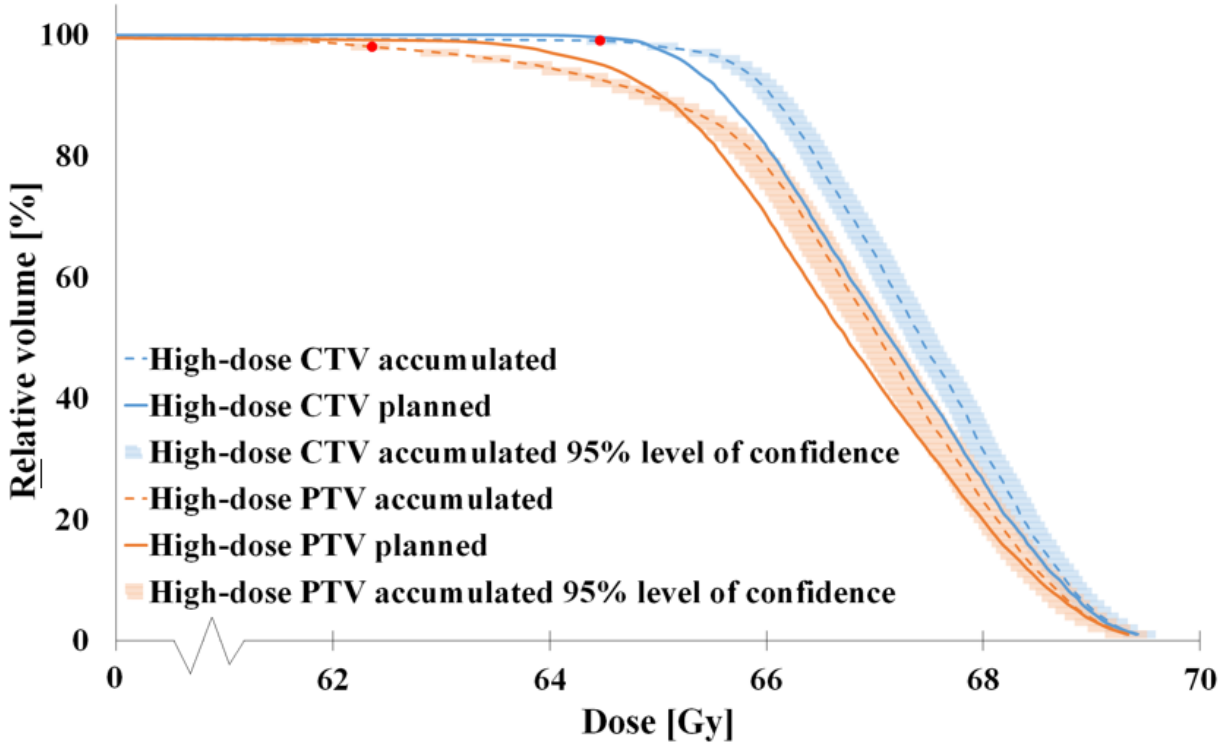


Figure 2: Dose-volume histogram (DVH) of the accumulated dose for patient 1's high-dose CTV and at the end of the 30 fraction treatment compared to the original planning DVHs. The accumulated $D_{98\%}$ and $D_{99\%}$ (red bullet points) of the high-dose PTV and CTV were significantly different at the 95% level of confidence from the planned values but still 97.8% and 99.5% of the planned $D_{98\%}$ and $D_{99\%}$, respectively.